Light-Induced Protein Conformational Changes Studied by Two-Dimensional Infrared and Transient Infrared Spectroscopies

---

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Claudio Zanobini
aus
Italien

Promotionskommission
Prof. Dr. Peter Hamm (Vorsitz)
Prof. Dr. Oliver Zerbe
Prof. Dr. Roland K.O. Sigel

Zürich, 2020

“Dedicated to my parents, brother and sister”

Contents

Abstract

Molecular spectroscopy, the branch of science which studies the interactions between light and matter, is an excellent tool to extract information on the system under investigation, where light is used as an external control element. The study of molecular vibrations employing infrared (IR) femtosecond pulses allows us to capture the molecular motions of the building blocks of life – proteins, in an ultrafast time-resolved manner. Here, two-dimensional infrared (2D-IR) and transient infrared (TRIR) spectroscopies have been used together to investigate the peptide-ligand recognition event and the associated allosteric transition in a PDZ2 domain. The recognition of the peptide-ligand interaction has been performed employing different site-specific azidohomoalanine (Aha) PDZ2 domain mutants, exploiting the sensitivity of this IR label to detect subtle changes of its chemical environment, without altering the chemical and biological properties of the studied system by its introduction. A peptide-ligand variant with an azobenzene moiety within its structure has been employed to promote a non-equilibrium allosteric transition in the PDZ2 domain by ultrafast photoisomerization of it. The experimental results reveal that a small number of discrete time scales are characteristic of the overall PDZ2 domain non-equilibrium allosteric transition. In addition, another biological system has been investigated, where the binding affinity between the two elements which constitute the enzyme RNase S can be controlled by light reversibly, achieving in one case complete unbinding.

Chapter 1

Introduction

Two-dimensional infrared (2D-IR) spectroscopy [1–8] is an innovative spectroscopic technique with a broad field of application, as it sheds light on the structure and dynamics of a multitude of systems. The complex dynamics of liquids [9–12], vibrational energy transfer mechanisms [13–16] and peptide-protein structural rearrangements [17–22] are all subjects which have been studied with it and represent only a small fraction of its full potential. Transient infrared (TRIR) spectroscopy is another useful tool to study the non-equilibrium dynamics of various processes where an intense UV or a visible (VIS) pulse is needed to promote a perturbation, whose effects on the system under investigation are probed using an ultrafast IR pulse [23–26]. The strength of these two spectroscopic techniques, 2D-IR and TRIR, is the combination of their intrinsic high time resolutions with their significant structural resolutions. These characteristics make 2D-IR and TRIR spectroscopies two complementary techniques to Nuclear Magnetic Resonance (NMR), where its inherent high structural resolution has been exploited from years for protein structure determinations [27–30], despite its intrinsic time resolution limitations. What determines the time resolution for the observation of a molecular process is the characteristic dephasing time of the transition under investigation, which is in the order of magnitude of a few picoseconds for standard molecular vibrations in solution, compared to milliseconds for typical nuclear spin transitions [31, 32]. The previous statement is valid in the homogeneous limit, or in other words, when the characteristic timescales of the molecular process in question are much faster than the timescale of the method employed for its observation (*e.g.* pulse duration, see Kubo line shape theory [33, 34]), which is often the case in NMR spectroscopy.

Both 2D-IR and TRIR spectroscopies allow us to observe the vibrational motions of the system under study, revealing information about changes of its structure and chemical environment. Depending on the type of molecular vibration which is investigated, a different kind of information can be obtained, observing the system from a distinct point of view, *e.g.* from a local or an overall perspective. Site-specific infrared labeling has been exploited as a strategy to monitor local structural changes in several biological systems [35–42]. The introduction of azido (-N_3), cyano (-CN), thiocyanate (-SCN), and thiol (-SH) functional groups into protein structures, have been employed by several researchers to address scientific questions of this type, where different interactions of the IR label with its chemical environment influence its spectral properties [35–44]. However, the observation of the system under study from an overall point of view can be beneficial, and the information obtained in this way is complementary to site-specific infrared labeling. As an example, the peptide bond is a well established infrared reporter concerning the secondary structure of proteins and peptides; the amide I mode, which is due mostly to the carbonyl (C=O) stretch vibration, is strongly sensitive to the backbone structural rearrangements of proteins and peptides [45]. A change of the secondary structure of such systems affects the amide I band, which can be used as an infrared reporter of the overall motions of proteins and peptides. It has to be stressed that, while a change of the secondary structure provokes a response of the amide I band, a change of the latter cannot easily be correlated to a structural rearrangement of the protein in a specific zone of it. Moreover, if the biological system under study is composed of more than one unit, as in the case of *e.g.* a protein-peptide or a protein-protein complex, it can be difficult to interpret the structural information contained in the amide I band. In short, the amide I mode gives insight into the overall structural rearrangements of the system and lacks site-specificity as anticipated previously, due to its intrinsically highly convoluted nature. The global information contained within the amide I band, which gives rise to congested infrared spectra, can be converted to site-specific information when isotope labeling strategy is employed [46–50]. In this case, the peptide bonds with a different isotopic composition, *e.g.* naturally abundant $^{12}C^{16}O$ versus $^{13}C^{16}O$ or $^{13}C^{18}O$ carbonyl groups, can be exploited to disentangle the contributions of the various peptide bonds to the overall amide I absorption band. It has been reported that the isotope labeled $^{13}C^{16}O$ and $^{13}C^{18}O$ carbonyl groups redshift the amide I absorption frequency of *circa* 20-35 and 60-70 cm^{-1} [46–48, 51], in comparison to the naturally abundant $^{12}C^{16}O$ case, respectively. Furthermore, the simultaneous employment of the various isotope labeled peptide bonds together with 2D-IR spectroscopy, can reveal crucial structural information of the biological system in question, due to the possible couplings and energy transfer processes which can be observed between the different vibrational oscillators as off-diagonal peaks in the 2D-IR spectrum [52].

The study of protein-peptide complexes is of paramount importance in science to understand binding mechanisms and molecular interactions, which *e.g.* can give rise to allosteric signal propagation and can be involved in cell signaling [53]. The coordination of multiple cell actions, which is crucial for the communication processes among cells and protein complexes, is at the basis of the regulation mechanisms in living organisms. This transmission of flow of information, which is referred in the literature as signal transduction [54], is composed by a series of cascading events where each of them often involves an allosteric protein complex [55, 56]. Allostery, *i.e.* the intriguing phenomenon of communication between two distant sites in a biological system, transmits specific signals through single proteins, large multidomain protein complexes, and long-range distances in cells to deliver precise information which induces a specific cellular response. Cell signaling involves allosteric propagation and PDZ domains are one example of single domain allosteric proteins which are present in several multidomain scaffold proteins [57, 58]. The allosteric response of PDZ domains induced upon ligand binding or upon ligand phosphorylation, have been investigated employing NMR spin relaxation experiments [59–62]. The binding event or the modulation of the binding affinity *via* ligand phosphorylation provoke variations of the protein side-chain fluctuation dynamics, which provide us with information about the phenomenon of allostery in PDZ domains [59–65]. It has to be stressed that NMR spin relaxation experiments give insight into the fluctuation dynamics of the molecular system under study at the chemical equilibrium, observing the relaxation dynamics of different active nuclear spin nuclei [66, 67]. In order to investigate the non-equilibrium allosteric signal propagation in PDZ domains, a spectroscopic technique with an intrinsic higher time resolution than NMR spectroscopy is needed, *e.g.* transient infrared (TRIR) spectroscopy. In addition to it, a non-equilibrium trigger is necessary to initialize the perturbation, ideally in a reversible manner. Azobenzene compounds are chemical tools which are ideal candidates for this task; they can be controlled by light and photoisomerized in both *trans* to *cis* and *cis* to *trans* directions employing two different wavelengths, reversibly [68, 69]. Moreover, the photoisomerization event happens on a ps time scale [70–72], which is compatible with the possibility to perform ultrafast time-resolved experiments. The power of photocontrol is the capability to use azobenzene compounds as optical switches, *i.e.* photoswitches which can be employed to modulate *e.g.* protein-peptide interactions and the functions of biomolecules, using light as a reversible external control element [73]. A relevant enzyme whose catalytic activity depends on protein-peptide interactions is the RNase S complex [74, 75]; it is composed of two units, a protein part called S-protein and a peptide moiety named S-peptide. What is remarkable of this enzyme is the fact that the crucial amino acids of it, which are part of the active site and are involved in the catalysis, are split into the two units, S-protein and S-peptide. In other words, the RNase S enzyme is not active when one of the two components is absent. It would be fascinating to photocontrol the interactions between S-protein and S-peptide and consequently to modulate the RNase S enzymatic activity by light, which can be an arduous task to be realized but definitely, with outstanding potential.

In the next sections of this chapter, the elements which have been used as molecular systems during my PhD activity will be introduced separately, one at the time, in the following order:

- PDZ domains
- Azidohomoalanine (Aha)
- Azobenzene photoswitches
- RNase A/RNase S

1.1 PDZ Domains

PDZ domains are protein-protein interaction modules whose abundance and diversity make them relevant to many cellular and biological processes. Several PDZ mediated interactions are involved in signaling transduction complexes [57–59, 76–85]. Different sorts of multidomain proteins, which contain PDZ modules within their structures, are shown in Fig. 1.1. Among the membrane-associated guanylate kinases (MAGUKs), postsynaptic density protein 95 (PSD-95) is one of the first three proteins where PDZ domains were identified. The other two, which give rise together with PSD-95 to the acronym "PDZ", are *Drosophila* disc large tumor suppressor (DLG1) and zonula occludens-1 (ZO-1) [86]. PDZ domains are mainly found in metazoans, especially in all three metazoan genomes: *Caenorhabditis elegans*, *D. melanogaster* and *Homo sapiens* [57, 87, 88]. Besides worms, flies, and mammals, PDZ domain homologues have been later discovered in yeast, bacteria, and plants [89, 90], highlighting how wide is the spread of this class of proteins in nature. PDZ domains are relatively small proteins of *circa* 80-100 amino acids, with a canonical fold of 6 β-strands (βA-βF) and 2 α-helices (αA-αB) (see Fig. 1.2A) [58, 59, 77, 92]. An extended groove between the strand βB and the helix αB forms a pocket which allows the binding of peptide-ligands *via* β-strand addition [93, 94]. The structural changes of PDZ domains upon ligand binding are minimal, *e.g.* PSD-95 reveals an α carbon root-mean-square deviation (RMSD) variation of 0.9 Å [57, 91]. The recognition of C-terminal binding sequences of *circa* 5 amino acid residues happens through the "carboxylate-binding loop" (see Fig. 1.2B), which has the following well-conserved sequence motif among the PDZ domains: R/K-X-X-X-G-L-G-F where X can be any amino acid [57]. It can be noted from its structure in Fig. 1.2B that many hydrogen bonds and some hydrophobic interactions are responsible for the recognition mechanism. The C-terminal target sequence can belong to a short peptide or to the extreme C-terminus of another PDZ domain [95].

The nomenclature for the target sequence is the following: the C-terminal residue

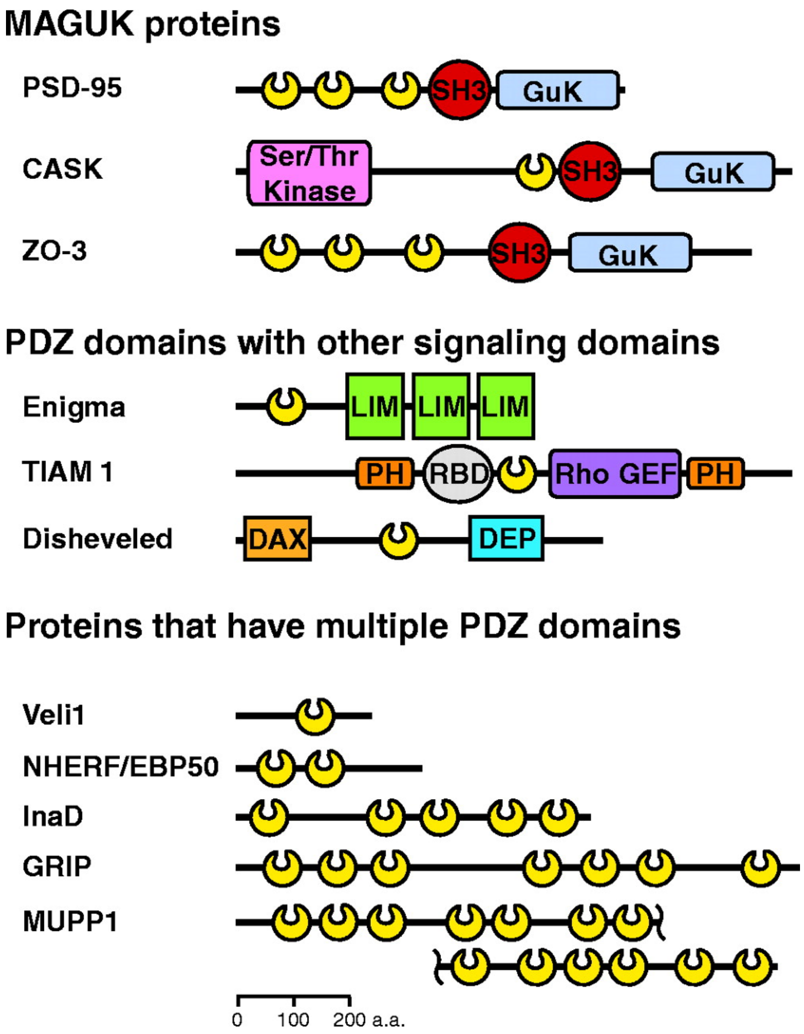


Figure 1.1: Examples of higher-order organization of PDZ domains found in signaling proteins. Proteins are indicated in black lines scaled to the length of the primary sequence of the protein; PDZ domains are shown in yellow. Other domains are indicated as abbreviations (from SMART) (Schultz *et al.* [88]) as follows: SH3, Src homology 3 domain; GuK, guanylate-kinase-like domain; LIM, zinc-binding domain present in Lin-11, Isl-1, Mec-3; PH, pleckstrin-homology domain; RBD, Raf-like Ras-binding domain; RhoGEF, Rho-like GTP-exchange factor; DAX, Dishevelled- and axin-homology domain; DEP, Dishevelled-, Egl-10- and pleckstrin-homology domain.
Adapted with permission from Harris, B. Z. & Lim, W. A. Mechanism and role of PDZ domains in signaling complex assembly. *Journal of Cell Science* **114,** 3219–3231 (2001). Order detail ID: 71889075, Order License: 4580150840910, ISSN: 1477-9137, Publisher: The Company of Biologists Ltd, Confirmation Number: 11811604, Order Date: 05/01/2019.

is called the P_0 residue and the next amino acids towards the N-terminus are named P_{-1}, P_{-2}, P_{-3}, *etc.*. Studies of Songyang *et al.* [95] and Schultz *et al.* [96] indicate that the residues P_0 and P_{-2} are the most important ones for the recognition mechanism. Depending on the preferences of PDZ domains to bind specific residues at these two sites, they can be classified at least in three principal classes in the

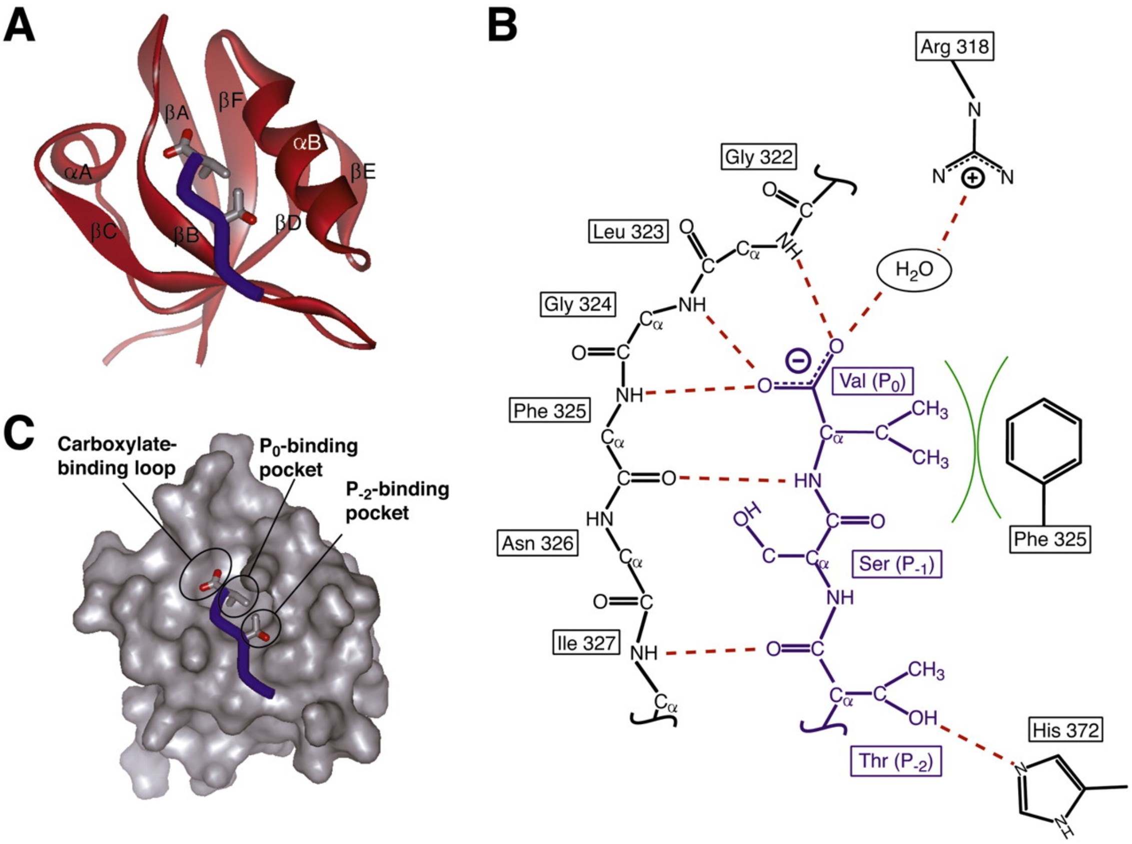


Figure 1.2: Structure of the PDZ domain and mechanism of peptide recognition. (A) Ribbon diagram of PSD-95 PDZ domain 3 (residues 306-394, shown in red) with a bound peptide (NH_2-KQTSV-COOH, shown in blue). Names of β-strands and α-helices are indicated. The side chains of the peptide P_0 residue (valine) and P_{-2} residue (threonine) are shown in stick form, as is the terminal carboxylate. (B) Diagram of the peptide-binding pocket. Residues in the PDZ-domain-binding pocket are shown in black; the peptide is shown in blue. Hydrogen bonds are drawn as red dotted lines, and hydrophobic packing is indicated by green arcs. (C) Solvent-accessible surface representation of the structure shown in (A) (probe radius = 1.4 Å). The peptide is drawn as in A, and key binding pockets are indicated by circles (Doyle *et al.* [91]).
Adapted with permission from Harris, B. Z. & Lim, W. A. Mechanism and role of PDZ domains in signaling complex assembly. *Journal of Cell Science* **114,** 3219–3231 (2001). Order detail ID: 71889075, Order License: 4580150840910, ISSN: 1477-9137, Publisher: The Company of Biologists Ltd, Confirmation Number: 11811604, Order Date: 05/01/2019.

following way [57, 95]:

- PDZ domain class I $\Rightarrow$ sequence motif recognition: S/T-X-Φ-COOH
- PDZ domain class II $\Rightarrow$ sequence motif recognition: Φ-X-Φ-COOH
- PDZ domain class III $\Rightarrow$ sequence motif recognition: X-X-C-COOH

where Φ are hydrophobic amino acids (usually valine (V), isoleucine (I) or leucine

(L)) and X can be any residue. As we can see from the classification, the P_0 residue is often hydrophobic, ensuring in this way a strong interaction with the inner hydrophobic part of the PDZ domain binding pocket. Regarding the PDZ domain class I, a histidine residue is present in a separate pocket (see Fig. 1.2B) allowing in this way a specific hydrogen bond interaction between the histidine N-3 nitrogen and the hydroxylated side chains of both serine (S) or threonine (T) residues present in the P_{-2} position of the ligand. The PDZ domain class II has the characteristic to have a leucine (L) or a methionine (M) instead of the above mentioned histidine, which allows the specific recognition of hydrophobic residues (Φ) at the P_{-2} position [97]. PDZ domains can bind selectively also internal peptide sequences, like in the case of α1-syntrophin PDZ domain (see Fig. 1.3); it can recognize the canonical C-terminal peptide sequence NH_2-VKESLV-COOH (Fig. 1.3A) [96] or the internal "β-finger" motif of the nNOS PDZ domain (Fig. 1.3B) [98]. This "double recognition" is not controversial but rather connected by the fulfillment of the same energetic requirements regarding the binding event. An internal motif can be suitable for the recognition when the overall tertiary structure of it, with its environment, mimics the conformation of the C-terminal target ligand. The nNos β-finger interacts with the syntrophin PDZ binding groove *via* a phenylalanine (F) and a threonine (T) residue at the P_0 and P_{-2} positions, where the C-terminus is substituted by a sharp β-turn.

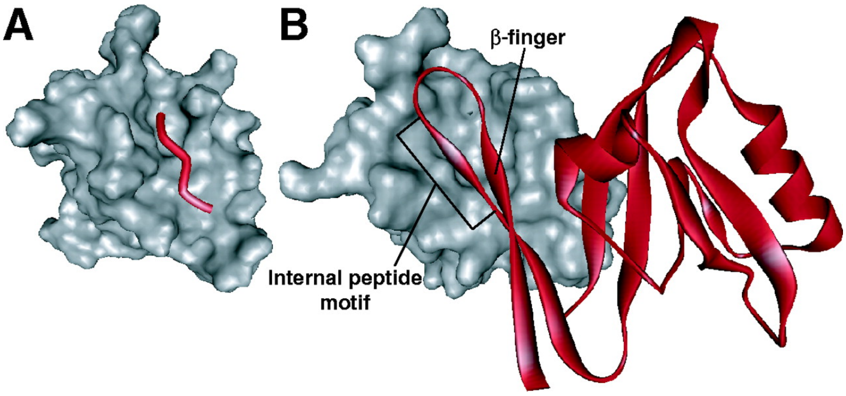


Figure 1.3: The same PDZ domain can recognize two structurally distinct ligands. The α1-syntrophin PDZ domain is shown as a gray solvent-accessible surface. (A) The ligand is canonical peptide NH_2-VKESLV-COOH, shown as a red tube (Schultz *et al.* [96]). (B) The ligand is the nNOS PDZ domain with its distinctive β-finger motif (indicated), shown as a red ribbon diagram. The internal peptide motif that mimics a C-terminal peptide is indicated (Hillier *et al.* [98]).
Adapted with permission from Harris, B. Z. & Lim, W. A. Mechanism and role of PDZ domains in signaling complex assembly. *Journal of Cell Science* **114,** 3219–3231 (2001). Order detail ID: 71889075, Order License: 4580150840910, ISSN: 1477-9137, Publisher: The Company of Biologists Ltd, Confirmation Number: 11811604, Order Date: 05/01/2019.

1.2 Azidohomoalanine (Aha)

Proteins are biomolecules which play a crucial role in life. They are constituted of a precise sequence of amino acids which allows them to fold into a characteristic three-dimensional structure – conformation to which a specific function is associated. Site-specific infrared labeling is an exceptional strategy to obtain selective information in biomolecules; the employment of infrared (IR) light to excite a specific molecular vibration can be used as a spectroscopic method to investigate the subtle changes of its local chemical environment in response to a specific event. In order to obtain site-specific insight on the regulatory mechanisms which affect several biological systems, originating from *e.g* protein-peptide interactions, the ideal IR label should have an absorption frequency in the spectral window between 1700 and 2800 cm^{-1}, where the vast absorption background of proteins is absent [35, 43, 44]. Moreover, it should be non-invasive in the sense that the insertion of the label should not affect the characteristics of the biological system in question. Azidohomoalanine (Aha) is a non-canonical amino acid (see Fig. 1.4) which absorbs at *circa* 2110 cm^{-1}, and it is employed as a sensitive, site-specific and minimally invasive IR label [35, 36, 99–103].

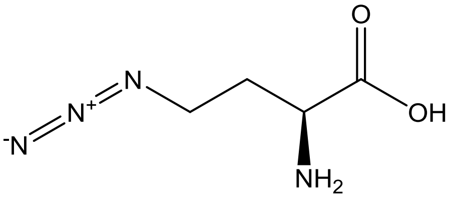


Figure 1.4: Azidohomoalanine (Aha)

Aha can be incorporated into proteins and peptides essentially in any position of their primary structures, using methionine auxotrophic protein expression strategy [99, 100] and fluorenylmethyloxycarbonyl (Fmoc) solid-phase peptide synthesis [104]. The azido group (-N_3), which is present in the structure of this unnatural amino acid, is a relatively strong infrared absorber with an extinction coefficient (ϵ) value of $\approx$ 350-400 M^{-1} cm^{-1} [102]. It has been shown that the variation of the polarity of the Aha chemical environment affects the absorption frequency of the vibrational label [35, 36, 99–103]. For example, Taskent-Sezgin *et al.* [36] observed a vibrational blue shift of the Aha label of $\approx$ 19 cm^{-1} in D_2O passing from a solvent-hindered/hydrophobic to a solvent-exposed/hydrophilic chemical environment, where the unfolding event of an NTL9 protein Aha mutant was promoted *via* thermal denaturation in this case. Furthermore, they demonstrated that the insertion of Aha into the NTL9 protein structure did not affect the folding of it, confirming the minimally invasive nature of the label.

1.3 Azobenzene Photoswitches

Azobenzene photoswitches are chemical tools which can be used as light-controlled triggers for biomolecules [68, 69, 105, 106]. In addition to these, caged compounds are another class of molecules which can be employed for the same purpose [107–109]. The latter act as inhibitors for biological systems through the presence of a "caging group" which makes the biomolecules inactive. Upon irradiation, the caging group can be cleaved, and the biological activity restored with the disadvantage that the light-induced liberation of the biomolecule is often an irreversible step. One of the advantages of employing azobenzene photoswitches relies on the reversibility of the isomerization process, which is ensured by their azobenzene moiety and its capability of *trans*, *cis* isomerization (see Fig. 1.5), employing two different wavelengths. Other two important aspects of azobenzene have to be underlined; firstly, the distance between the carbons at the *para* positions of the phenyl rings changes of ≈ 3.5 Å upon its isomerization [68]. Secondly, the photoinduced isomerization of azobenzene happens in the sub-ps regime [70, 71]. Both the previous aspects and the reversibility of the azobenzene isomerization process highlight the potential of this molecule to be used as an ultrafast and reversible phototrigger in chemistry [69].

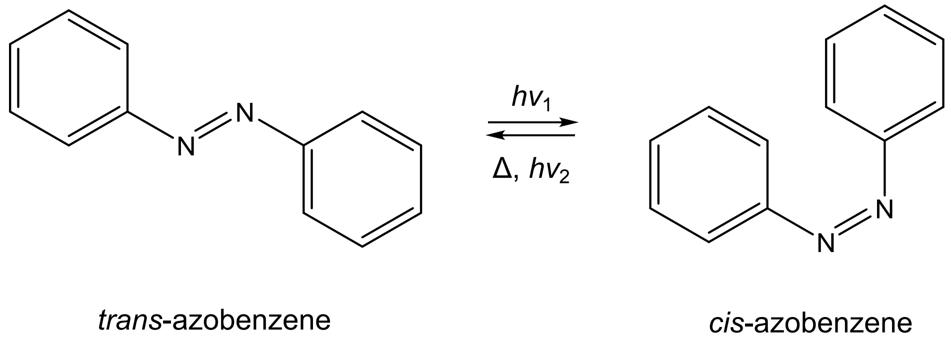


Figure 1.5: Azobenzene *trans* and *cis* isomers

Azobenzene compounds were discovered in the mid-1800s [110, 111] and since then, they were employed for several years as synthetic colouring agents to be used in the dye industry. Despite their recognition by the scientific community for more than a century, scientists started to study the photochemistry of azobenzene and of its derivative only in the mid-1900s. The *trans* and *cis* UV-VIS absorption spectra of azobenzene are distinctive of each isomeric form, where two main absorption bands are present and both their spectral positions and intensities are affected by the isomerization process [68]. The *trans* isomer is more stable than the *cis* form with the difference of *circa* 10-12 kcal mol^{-1} or in other words, the percentage of azobenzene molecules in the *trans* configuration is $\gg 99\%$, in the dark at the equilibrium [112]. The *trans* form has a strong absorption band at $\lambda_{max} \approx 320$ nm ($\epsilon \approx 22000$ M^{-1} cm^{-1}) and a much weaker one at $\lambda_{max} \approx 450$ nm ($\epsilon \approx 400$ M^{-1} cm^{-1}). The difference in the extinction coefficients is due to the intrinsic nature of the electronic transitions involved where the one in the UV region is symmetry allowed ($\pi \rightarrow \pi^*$ excitation), while the one in the visible is symmetry forbidden ($n \rightarrow \pi^*$ excitation) [113]. The corresponding $\pi \rightarrow \pi^*$ absorption band of the *cis* isomer is weaker and blue-shifted at $\lambda_{max} \approx 270$ nm ($\epsilon \approx 5000$ M^{-1} cm^{-1}), while the $n \rightarrow \pi^*$ band does not shift considerably but it becomes stronger (λ_{max} $\approx$ 440-450 nm, $\epsilon \approx 1500$ M^{-1} cm^{-1}).

Both quantum yields and mechanisms of the isomerization process depend on the excitation wavelengths employed to promote it, solvent polarity and viscosity, pressure and temperature [114–119]. The presence of substituents and their positions on the phenyl rings of the azobenzene moiety influence both properties too [120]. Regarding the quantum yields, the *cis* to *trans* isomerization is favoured respect to the *trans* to *cis* one for both $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions. Moreover, between the $\pi \rightarrow \pi^*$ and the $n \rightarrow \pi^*$ transitions, the quantum yields of the $n \rightarrow \pi^*$ transition are higher. For example, in the case of n-Hexane as solvent, the *cis* to *trans* isomerization quantum yields ($\Phi_{cis \rightarrow trans}$) of azobenzene are ≈ 0.44 and ≈ 0.56 as a result of the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ excitations, while the corresponding ones for the *trans* to *cis* process ($\Phi_{trans \rightarrow cis}$) are ≈ 0.11 and ≈ 0.25, respectively [113, 116, 121]. It has to be stressed that after illumination of azobenzene, the relative concentrations of the *trans* and *cis* forms at the photoequilibrium depend on the quantum yields for both $\Phi_{trans \rightarrow cis}$ and $\Phi_{cis \rightarrow trans}$ isomerization processes and on the relative extinction coefficient values of the *trans* and *cis* species; both the previous parameters are strongly dependent on the wavelength employed to conduct the experiment and while thermal relaxation gives rise to essentially 100% of the *trans* isomer, photoisomerization of azobenzene *via* $\pi \rightarrow \pi^*$ or $n \rightarrow \pi^*$ excitation leads to different amount of *trans* and *cis* isomers at the photoequilibrium. For example, *Fischer* reported that illumination of azobenzene in methylcyclopentane as solvent at 0°C with 313 nm or 436 nm radiation results in a photostationary state of $\approx 22\%$ and $\approx 88\%$ of *trans* azobenzene, respectively [115]. Concerning the isomerization mechanism, rotation, inversion, concerted inversion and inversion-assisted rotation have been proposed to clarify the complex photochemistry of azobenzene, which has been a topic of debate in the scientific community for several years [118, 119, 122–127].

The research group of *Andrew Woolley* at the University of Toronto performed a tremendous amount of work in the research field of photoswiches; a multitude of variants have been designed and synthesized, whose characteristics differ in terms of UV-VIS absorption spectra, isomerization wavelengths, steric hindrance, thermal *cis* to *trans* relaxation times and solubility in water [68, 128–134]. Various research groups (*Hamm*, *Woolley*, *Wachtveitl*, just to mention a few of them) took into account the idea of photoswitching peptides [71, 72, 135–144]. *Hamm* and *Wachtveitl* performed ultrafast time-resolved experiments in this regard [71, 72, 135, 136, 138–140]. *Woolley* and coworkers studied the effect of photoswitching on the helical content of peptides, varying the linking distance of the photoswitch within the peptide structure [142, 143]. As a result of it, it has been observed that the anchoring spacings *i+4* and *i+7* have the effect to increase the amount of the helical content upon *trans* to *cis* peptide photoswitching, as revealed by circular dichroism (CD) measurements [142]. On the other hand, spacing *i+11* stabilizes the peptide helix in the *trans* configuration of the photoswitch. Moreover, *Hamm* and collaborators explored the research field of protein photoswitching, carrying out ultrafast time-resolved experiments on a photoswitchable PDZ2 domain protein [145, 146].

In order to perform ultrafast time-resolved experiments on biomolecules, the following requirements have to be taken into account:

- The photoswitch has to be reactive towards the target molecule to ensure the chemical linkage between the two units. The specificity of the chemical bond is guaranteed by the reactivity of the two free thiol groups of the cysteine (C) residues on the biomolecule with the two acyl halide functional groups on the photoswitch [131, 134, 147].

- The photoswitch should be as much as possible soluble in water for two reasons: firstly, the linking reaction should be performed in water, the natural solvent of biomolecules. Secondly, the photoswitch should decrease the solubility of the biomolecule in water as little as possible, once the latter is linked to it. High concentration values in the mM regime are needed to perform ultrafast time-resolved infrared experiments.

- Ideally, the photoswitch should lock the biomolecule in an "off" functional state when the azobenzene moiety is in one of the two *cis* or *trans* configurations, while the biomolecule should be active when the chromophore is in the other isomerization state. Depending on the positions where the photoswitch is linked to the biomolecule, the extent to which the previous statement is achieved can change dramatically.

Bearing in mind the previous considerations, two different photoswitch variants (see Fig. 1.6) have been employed by *Hamm* and coworkers; both photoswitch versions, which differ mostly for their solubility in water [131, 134], are present in the articles [35, 101, 148] and chapters of this thesis, as well as in other papers published in the past by previous *Hamm*'s group members [72, 135, 145, 146, 149]. The water-soluble photoswitch has two hydrophilic sulfonate groups (—SO_3^-) at the *meta* positions of the azobenzene moiety which ensure its water solubility. Both water-soluble and water-insoluble photoswitch variants have comparable UV-VIS absorption spectra and extinction coefficients [131].

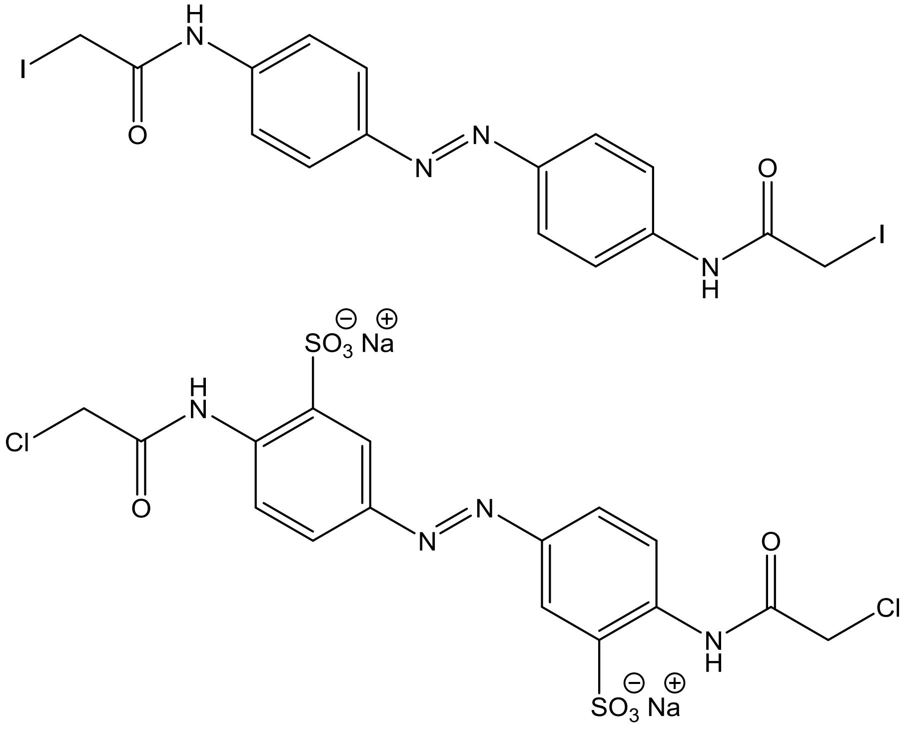


Figure 1.6: Water-insoluble and water-soluble photoswitches (top and bottom chemical structures, respectively).

The water-soluble version in Fig. 1.6 (bottom structure) has been used to modulate the conformation of a PDZ2 domain and of the S-peptide unit in the RNase S complex by photoisomerization of it (see chapters 2 and 5 for details). The water-insoluble variant (see Fig. 1.6, top structure) has been employed to obtain a novel peptide-ligand for the same PDZ2 domain as before (see chapters 3 and 4 for details). The water-insoluble photoswitch showed a higher reactivity in the linking reaction with the latter peptide-ligand, in comparison to the water-soluble version, and this is the reason for its choice and usage in this case. Due to the presence of hydrophilic amino acid residues in the primary structure of this peptide-ligand, the linking of the water-insoluble photoswitch to it did not compromise the water-solubility of the linking product greatly, which is water-soluble at mM concentrations. However, it is recommended to use the water-soluble version with the sulfonate groups (—SO_3^-) whenever it is possible.

1.4 RNase A/RNase S

Ribonuclease A (RNase A) is an enzyme which does not need presentations; it has been the reference point for years in the field of protein stability, folding, and chemistry as well as in enzymology and molecular evolution [75, 150, 151]. Two Nobel Prizes in Chemistry were assigned for the work on this enzyme; *Christian B. Anfinsen* was honoured in 1972 "for his work on ribonuclease, especially concerning the connection between the amino acid sequence and the biologically active conformation", while *Stanford Moore*/*William H. Stein* "for their contribution to the understanding of the connection between chemical structure and catalytic activity of the active centre of the ribonuclease molecule" (from The Royal Swedish Academy of Sciences, https://www.nobelprize.org/prizes/chemistry/1972/summary/) [152–154]. A few years after, *Bruce Merrifield* was rewarded in 1984 "for his development of methodology for chemical synthesis on a solid matrix" (from The Royal Swedish Academy of Sciences, https://www.nobelprize.org/prizes/chemistry/1984/summary/) with which the total synthesis of RNase A was performed [155, 156].

As the word suggests, ribonucleases are enzymes which catalyze the depolymerization of ribonucleic acid (RNA) [157, 158]. The bovine pancreatic RNase A was discovered in the pancreas of ruminants, where large quantities of RNA needed to be digested. In this way, the phosphorus and nitrogen contained in it could be recovered and recycled by the organism of ruminants in order to satisfy their metabolic needs. It is interesting to note that the amount of pancreatic RNase A in vertebrates changes abruptly from species to species, for example in a cow, kangaroo, mouse, and horse there are $\approx$ 1200, 600, 400 and 25 times the equivalent amount of RNase A contained in the pancreas of a human, respectively [159].

Isolation methods of RNase A from bovine pancreas were proposed by *Kunitz* and *McDonald* in the mid-1900s [160–162]. The drastic conditions employed for the RNase A purification steps (*e.g.* storage solution 0.25 N sulfuric acid at 0-5°C followed by heating at 95-100°C for 5 minutes at pH 3) highlight the incredible stability of this enzyme.

The four disulfide bridges among the cysteine (C) residues C26-C84, C58-C110, C40-C95, C65-C72 (see Fig. 1.7, amino acids coloured in yellow) contribute to the overall conformational stability of RNase A. The C26-C84 and C58-C110 disulfide bonds establish a chemical connection between α-helical and β-sheet regions of the enzyme, enhancing stability more than do the C40-C95 and C65-C72 disulfide linkages, which connect different surface loop regions of RNase A [75, 147]. The C40-C95 and C65-C72 disulfide bridges contribute also to the efficiency of the enzymatic activity, allowing the proper positioning of lysines K41 and K66 in the active site of RNase A [163]. All the four disulfide bonds are involved in the folding mechanism of the enzyme as well as the *cis* oriented proline residues (P93 and P114, highlighted in grey in Fig. 1.7) [75, 164, 165].

The amino acids which are crucial for the catalytic activity of RNase A are H12, H119, and K41 (see Fig. 1.7, residues highlighted in red). The imidazole groups of the two histidines act one as a base (H12) and the other one as an acid (H119) in the catalytic mechanisms, while K41 stabilizes the negatively charged transition state during the RNA breakage [75, 166–168]. Other residues like Q11, F120 and D121 are involved in the stabilization of the transition state too, while positively charged amino acids such as K7 and K66 and R10 play a key role in the recognition of the negatively charged phosphate backbone of RNA [75, 169].

Moreover, the cleavage of the specific peptide bond between the 20th and the 21st amino acids is observed when RNase A is treated with the enzyme subtilisin, giving rise to two subunits named S-peptide (residues 1-20) and S-protein (residues 21-124) [170]. The two components together form the adduct RNase S, where the letter “S” refers to the subtilisin protease. It is interesting to note that neither S-protein or S-peptide are enzymatically active, while the adduct RNase S has a catalytic activity comparable to the intact enzyme RNase A [75, 150, 170]. The first 20 amino acids of the RNase A primary structure have been coloured in blue in Fig. 1.7, in order to highlight the “S-peptide part” of the enzyme which assumes an α-helical conformation.

Despite the fact that the tertiary structures of RNase A and of the complex RNase S are similar, S-peptide assumes a disordered and an unfolded structure when it is unbound from S-protein, while S-protein adopts a much more flexible conformation in the absence of S-peptide [74, 171–174]. These characteristics make RNase S an interesting model system to be studied in order to clarify the issue concerning the binding mechanism between S-protein and S-peptide, which is not well understood yet [175–177]. Two association mechanisms have been proposed as limiting binding scenarios with the names “induced fit” and “conformational selection” [178], where the discrimination between the two is based on the chronological order of the binding and conformational change events. The former hypothesis considers the binding of S-protein with unfolded S-peptide, which is followed by the S-peptide folding. The second mechanism assumes the S-protein recognition of a small fraction of free already folded S-peptide molecules into an α-helical conformation, instead. Since both S-protein and S-peptide undergo disorder to order transition upon binding, the previous approach seems to be too simplistic for the description of the RNase S formation mechanism as proposed by Dogan *et al.* [179], where multi-step binding reactions have to be taken into account as regards to complex molecular systems like intrinsically disordered proteins.

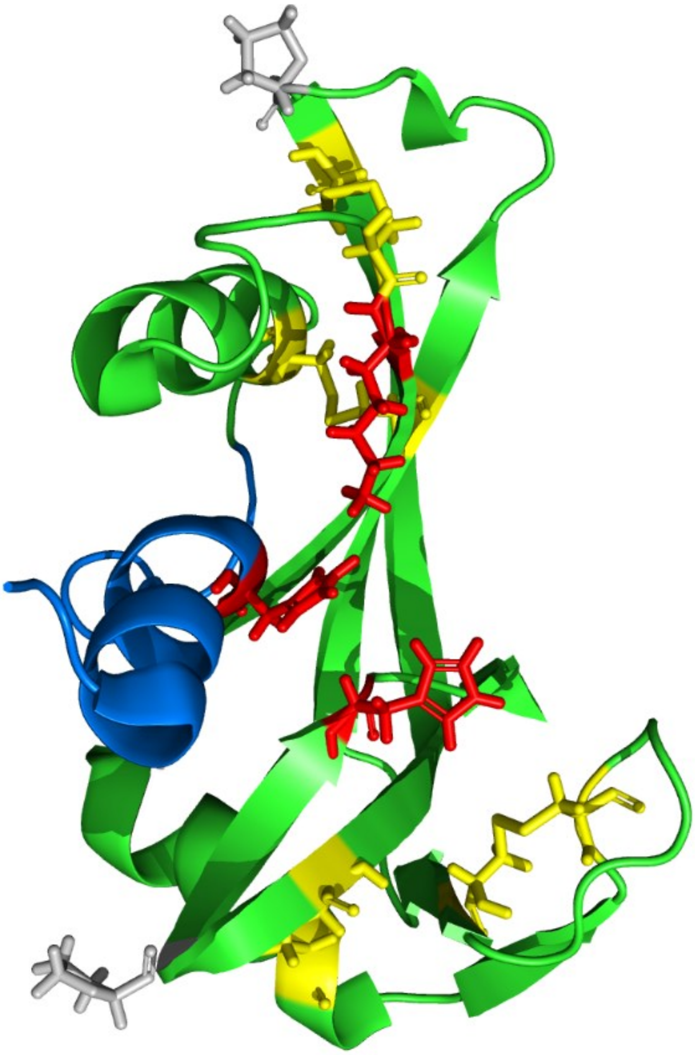


Figure 1.7: Ribonuclease A (RNase A). The picture was processed with the software PyMOL using the PDB structure 5RSA [180]. Different colours were used in order to highlight distinct aspects of the enzyme. S-peptide part (blue), S-protein part (green), disulfide bridges formed by the cysteine residues (yellow), crucial amino acids for the catalysis (H12, H119 and K41, red) and *cis* oriented prolines (P93 and P114, grey).

Chapter 2

2D-IR Spectroscopy of an Aha Labeled Photoswitchable PDZ2 Domain

---

Here, the study of a PDZ2 domain system is presented where the site-specific infrared label azidohomoalanine (Aha) [35, 36, 99–103] has been exploited as a sensitive, local infrared reporter for the PDZ2 domain chemical environment by the use of 2D-IR spectroscopy [101]. Aha has been inserted in six different positions (one at the time) within the protein structure to verify if the structural changes between the apo and holo states of the PDZ2 domain (with ligand unbound and bound, respectively) could have been sensed by Aha to obtain site-specific information about the overall apo/holo process. The ligand-binding has been mimicked (it has been confirmed by NMR spectroscopy) by photoisomerization of an azobenzene moiety covalently linked to the binding groove of the PDZ2 domain, where the *cis* and *trans* configurations of the photoswitch [131, 141, 143] mimicked the apo and holo states of the PDZ2 domain [145]. The photoswitch which has been used in this case is the water-soluble version reported in Fig. 1.6 (bottom chemical structure) [131, 141, 143]. The 2D-IR responses of the six different Aha mutants were recorded with the *trans* and *cis* photoswitch configurations. Upon photoswitching, an interesting and uncommon result was observed for one particular mutant, as described in the following paper.

- ∗ I contributed to this paper confirming the results in Fig. 3 and Fig. S3 obtained for the mutants L78Aha and N80Aha by the acquisition of their purely absorptive 2D-IR difference spectra, with the proof of the L78Aha stability before and upon illumination of it with a 370 nm cw diode laser (Fig. S4), and confirming the results reported with the FTIR difference spectra in Fig. S5.

THE JOURNAL OF PHYSICAL CHEMISTRY A
Article
Cite This: J. Phys. Chem. A 2017, 121, 9435−9445
pubs.acs.org/JPCA

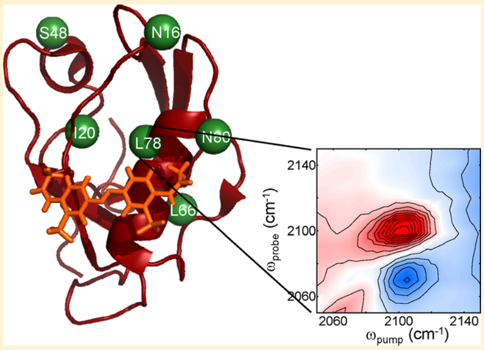
2D-IR Spectroscopy of an AHA Labeled Photoswitchable PDZ2 Domain

Published as part of The Journal of Physical Chemistry virtual special issue "Time-Resolved Vibrational Spectroscopy".

Brigitte Stucki-Buchli, Philip J. M. Johnson, Olga Bozovic, Claudio Zanobini, Klemens L. Koziol, Peter Hamm,* Adnan Gulzar, Steffen Wolf, Sebastian Buchenberg, and Gerhard Stock*

Department of Chemistry, University of Zurich, 8075 Zurich, Switzerland
Biomolecular Dynamics, Institute of Physics, Albert Ludwigs University, 79104 Freiburg, Germany

Supporting Information



ABSTRACT: We explore the capability of the non-natural amino acid azidohomoalanine (AHA) as an IR label to sense relatively small structural changes in proteins with the help of 2D IR difference spectroscopy. To that end, we AHA-labeled an allosteric protein (the PDZ2 domain from human tyrosine-phosphatase 1E) and furthermore covalently linked it to an azobenzene-derived photoswitch as to mimic its conformational transition upon ligand binding. To determine the strengths and limitations of the AHA label, in total six mutants have been investigated with the label at sites with varying properties. Only one mutant revealed a measurable 2D IR difference signal. In contrast to the commonly observed frequency shifts that report on the degree of solvation, in this case we observe an *intensity* change. To understand this spectral response, we performed classical MD simulations, evaluating local contacts of the AHA labels to water molecules and protein side chains and calculating the vibrational frequency on the basis of an electrostatic model. Although these simulations revealed in part significant and complex changes of the number of intraprotein and water contacts upon *trans*−*cis* photoisomerization, they could not provide a clear explanation of why this one label would stick out. Subsequent quantum-chemistry calculations suggest that the response is the result of an electronic interaction involving charge transfer of the azido group with sulfonate groups from the photoswitch. To the best of our knowledge, such an effect has not been described before.

I. INTRODUCTION

Proteins are dynamical objects. The structural dynamics of proteins involve equilibrium processes, such as thermally driven fluctuation, as well as nonequilibrium processes, such as the conformational transition in a light-triggered protein. Vibrational spectroscopy provides an inherent picosecond time resolution to study both equilibrium and nonequilibrium processes. However, obtaining site-selective information from vibrational spectroscopy will in general require vibrational labels, because the IR spectrum of a molecule of that size is no longer resolved into its normal modes. Ideally, such vibrational labels should absorb outside of the congested region of the absorption spectrum of a protein[1−5] to discriminate it from a huge background. Various distinct molecular groups have been suggested for that purpose: −SH vibrations of cysteines,[6] −CD vibrations of deuterated amino acids,[7−9] −C≡O vibrations of metal−carbonyls either from natural cofactors such as a heme group[10−12] or from complexes that can be bound to amino acids in a post-translational step,[13−15] as well as $-N_3$,[16−22] −C≡N,[23,24] and −SCN[25] vibrations from non-natural amino acids. All these molecular groups have in common that they absorb in a spectral window between ≈1700 and ≈2800 cm^{-1}, where essentially no fundamental modes of natural proteins are found (with the one exception of the −SH vibrations of cysteines), and where the only background originates from the still quite strong but very broad and featureless absorption of the solvent water.

In addition to having a frequency in that spectral window, a good label should fulfill the following criteria:

- The extinction coefficient of the label should be large enough that it can be detected at reasonable concentrations. As most proteins are not soluble at high concentrations, the goal is to measure at concentrations of around 1 mM or below. This concentration range is comparable to what is commonly used for NMR measurements.
- It should be versatile and incorporable at essentially any position of a protein with good yields and at a high purity.

Received: September 29, 2017
Revised: November 20, 2017
Published: November 21, 2017

DOI: 10.1021/acs.jpca.7b09675
J. Phys. Chem. A 2017, 121, 9435−9445

- It should not significantly perturb the structure and the stability of a protein.
- The label should be sensitive to its environment, e.g., sense the polarity or hydrophobicity of its surrounding.

With these criteria in mind, we currently concentrate on the non-natural amino acid azidohomoalanine (AHA) as a label, which contains an azido group ($-N_3$) that absorbs at around 2100 cm^{-1}. It has a reasonably high extinction coefficient of 300−400 M^{-1} cm^{-1},[20] large enough to be measured at concentrations well in the submillimolar regime by 2D IR spectroscopy.[4,22] As a small amino acid that is a methionine analog, it can be incorporated also into larger proteins (which can no longer be synthesized on a peptide synthesizer) at essentially any position by a methionine auxotrophic protein expression strategy.[26,27] It does not perturb protein properties very much, as evidenced, for example, by the fact that labeling a peptide ligand with AHA affects its binding affinity to a PDZ2 domain only to a small extent.[20] Finally, AHA has been shown to be a sensitive probe of its environment. For example, when AHA is buried in the hydrophobic core in the folded state,[17] the azido band blue-shifts by up to 20 cm^{-1} upon protein unfolding. Along the same lines, it can also be used to detect binding of an AHA-labeled peptide ligand to a larger protein, in which case the degree of solvation of the label diminishes, causing a red shift of its vibration.[20,22] Such shifts have been explained mainly by changes in hydrogen bonding to the first and/or third nitrogen atom of the azido group as well as to variations in the angle between the hydrogen bond and the azido group.[21,28] Furthermore, the vibrational frequency is also influenced by the polarity of the environment;[21] e.g., Coulomb interactions to the middle nitrogen atom of the azido group can shift the vibration.[18]

In the present study, we explore the capability of the AHA label to sense relatively small structural changes of a protein, i.e., changes much smaller than those occurring upon unfolding. To that end, we employ a protein construct that we have designed and characterized recently, Figure 1.[29−32] That is, we have chosen the second PDZ (PDZ2) domain from human tyrosine-phosphatase 1E (hPTP1E), which has been studied extensively as a model allosteric protein from different perspectives, i.e., structural,[33−35] dynamical,[36−38] or computational.[31,32,39−48] To investigate the allosteric mechanism by transient IR spectroscopy, we have covalently linked an azobenzene derivative across the binding groove of the PDZ2 domain in such a way that the light-driven *trans*−*cis* isomerization of the photoswitch induces a structural transition in the protein, which mimics ligand unbinding in the native system.[29] By NMR analysis (PDB entries: 2M0Z and 2M10), we have confirmed that the structural changes, which are of the order of 1 Å, are of similar size as in the native system.

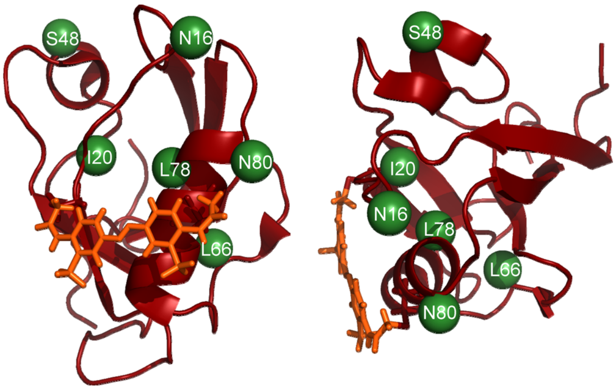


Figure 1. Photoswitchable PDZ2 domain with the positions of amino acids, which have been mutated to AHA, indicated in green. Only one amino acid has been mutated at a time in the experiment. Left: front view to the binding site of the protein. Right: top view.

We have shown by transient IR spectroscopy that the protein responds to the light-induced perturbation on two time scales, one extending up to ≈100 ns that reflects the opening of the binding groove (which we know because luckily we could isolate one specific normal mode localized on the azobenzene derivative), and a second phase on a 10 μs time scale.[29] We have no information from experiment on the structural nature of the second phase, because we did not employ any label that would reveal site-selective information. On the basis of molecular dynamics (MD) simulations, we suggested that it involves some of the more flexible and remote loop regions of the protein and/or the termini.[31,32] MD simulations also suggested[29] that the water solvation shell changes on the time scale of the binding groove opening even relatively far away from the perturbation, and we proposed that both aspects (i.e., structural changes of regions far away from the binding groove and/or changes in solvation) might be possible mechanisms of allostery. On the basis of the experience from previous works employing AHA in other molecular systems,[17−22,28] it appears possible that this label can sense such effects in a photoswitchable PDZ2 domain, and it is the goal of the present paper to explore whether this is indeed the case.

As a first step in this direction, we consider stationary difference spectra comparing the two states of the protein, and leave time-dependent transient experiments for a future publication. To explore the capability of the AHA label as a probe of the change of the local structure, the following criteria have been used to select the positions for the AHA label:

- Amino acids close (L78, N80, I20) versus far away (N16, S48, L66) from the photoswitch have been selected, to investigate the dependence on the distance from the perturbation.
- Amino acids within secondary structural elements (L78, S48) versus loop regions (N16, L66) have been selected. Whereas the PDZ2 domain undergoes only small shifts in the more rigid secondary structural elements, larger conformational changes can be observed in the flexible loops.[29]
- Surface exposed amino acids (N80, N16, S48) versus amino acids buried in the hydrophobic core (L66, I20) have been selected. Only the former will sense changes in protein solvation.
- Finally, we considered only amino acids of similar size for mutation, and only neutral amino acids to avoid shifts of the isoelectric point of the protein.

Parts a and b of Figure 1 show all amino acids that have been mutated to AHA (for each mutation, only one amino acid at a time). We use 2D IR spectroscopy instead of FTIR spectroscopy owing to the inherent sensitivity gain of 2D IR, which for the most part originates from its quadratic dependence on the extinction coefficient that significantly reduces the solvent background in a relative sense.[4] Furthermore, because the measurement beams are small, and because we can light-induce the *trans*-to-*cis* transition, extraordinary low amounts of protein sample are needed (≲1 nmol) in 2D IR difference spectroscopy. These low amounts compensate for the fact that protein preparation is quite tedious.

II. MATERIAL AND METHODS

II.A. AHA-Labeled, Photoswitchable PDZ2 Domain. To produce protein containing the non-natural amino acid AHA, we have modifed a protocol that has been described before[27] as follows: Starting from the previously used pET30a(+) vector containing two mutations S21C and E76C for the cross-linking of the photoswitch,[29] additional single residues were mutated to methionine for the insertion of AHA by site-directed mutagenesis (QuikChange; Agilent, Santa Clara, CA). A methionine auxotrophic cell strain (*E. coli* B834(DE3); Novagen, Merck Millipore, Darmstadt, Germany) was used for the protein expression. Cell cultures were grown in LB medium with 30 μg/mL kanamycin at 37 °C to an OD600 of 0.8. The LB cultures were centrifuged at 3500 g and 20 °C for 20 min and afterward immediately resuspended in a minimal medium (SelenoMet Base medium with SelenoMet Nutrient mix from Molecular Dimensions, Newmarket, U.K.) supplemented with 100 μg/mL AHA (Bapeks, Riga, LV) and 30 μg/mL kanamycin, where they were incubated at 37 °C for 30 min to use up residual methionine. IPTG (1 mM) was added and the protein was expressed at 37 °C. AHA is a potentially reactive amino acid and can be modified during protein expression and purifcation. A short expression time of 4 h was therefore chosen to minimize protein modifications, despite a somewhat lower yield. The cell cultures were centrifuged at 3500 g and 4 °C for 20 min and then stored at −20 °C.

The His-tagged protein was purified using nickel magnetic beads (Bimake.com, Houston, TX, USA) following the recommended protocol. Tris buffer (50 mM), pH 8.5, was used for cell lysis; the protein was denatured with 6 M guanidinium chloride and then refolded on the nickel magnetic beads. Residual nickel, which was bound to the protein, was removed by adding 40 mM EDTA and incubating overnight at 4 °C. A yield of about 7 mg protein per liter of cell culture was determined using a Bradford assay. SDS-PAGE and mass spectrometry were performed to control the purity of the protein.

As photoswitch, we employed an azobenzene derivative containing two sulfonate ($-SO^{3-}$) groups to increase its solubility in water.[49] For its cross-linking, a protocol similar to the one described in ref 29 was used. However, the reduction of the cysteins, to which the photoswitch binds, had to be performed under milder conditions, because any reducing agent that is commonly used for that purpose would also reduce the azido group of the AHA, resulting in a primary amine. The reaction conditions had to optimized, so that most of the disulfide bridges were reduced while most of the AHA remained intact. That is, 1 mM TCEP (from a 100 mM TCEP stock at pH 8.5) was added to the protein at a concentration of around 50−200 μM in 50 mM Tris buffer, pH 8.5, 500 mM imidazole and 40 mM EDTA. This solution was incubated at room temperature for no longer than 15 min. In principle, the reduction of disulfide bonds could be omitted completely to avoid any destruction of AHA, which, however, would lower the yield of the cross-linking reaction dramatically.

As in ref 29, the reducing agent was subsequently removed by desalting chromatography (HiPrep column, GE Healthcare, 50 mM Tris buffer, pH 8.5), and the protein was cross-linked to the photoswitch under an oxygen free (nitrogen) atmosphere at room temperature for at least 6 h. Cross-linking was performed in a highly diluted solution to minimize the formation of oligomers (10 μM protein to 100 μM photoswitch). The cross-linked protein was purified using anion exchange chromatography (HiTrapQ column, GE Healthcare). The His-tag was removed by digestion with HRV 3C protease, and the cleaved protein was purifed with nickel affinity chromatography (His-Trap HP column, GE Healthcare). The purified protein was concentrated and desalted into 50 mM borate buffer with 150 mM NaCl at pH 8.5. Finally, the protein was lyophilized and dissolved in D_2O. Mass spectra of all mutants considered in this study are shown in Figure S1 (Supporting Information), emphasizing the excellent purity of the final labeled and cross-linked protein.

II.B. Difference 2D IR Spectroscopy. For the difference 2D IR spectroscopy, we used an instrument described before.[20,22] In brief, mid-IR pulses were generated in a home-built two stage OPA with a difference mixing stage[50] pumped by a commercial Ti:S amplified laser system (Spitfire, Spectra Physics) running at 5 kHz. The OPA yielded pulses at 4.7 μm and ≈3 μJ per pulse with an energy stability better than 0.3% at 500 shots. The 2D-IR instrument used a four wave mixing phase-matching geometry employing a HeNe trace beam to accurately determine the delay times[51] and a polarization-based balanced heterodyne detection.[52] The signal was detected on a 2 × 32 MCT detector array after dispersing it in a spectrograph with a resolution of 7 cm^{-1}. A photoelastic modulator (PEM) was used to induce a quasi-phase shift on pulses 1 and 2 to suppress scattering.[53] The time domain data were collected into 2.11 fs long time bins (defined by the HeNe wavelength) with a maximum scanning time of 3 ps, revealing a spectral resolution of 2.7 cm^{-1} after zero-padding by a factor 2 and subsequent Fourier transformation. Purely absorptive spectra were obtained by alternative scanning of beams 1 and 2 backward in time. The population time was kept constant at 300 fs to minimize nonresonant effects from overlapping excitation pulses.

At the protein concentrations considered in this study (1 mM or below), the AHA signal is buried under the background from the D_2O buffer (essentially the wing of the OD-stretch vibration that is centered at ~2500 cm^{-1}). We therefore measured difference 2D IR spectra; i.e., we first measured a spectrum of the dark adapted protein, in which case the photoswitch is in the *trans* configuration. The sample was then switched into *cis* by illuminating it at 370 nm from a cw-diode laser (CrystaLaser CL-2000) for about 3 min, and a second 2D IR spectrum was measured without changing any of the alignment (but with continuous illumination with the 370 nm laser). The difference of the two spectra was then calculated after phasing them independently, using the phase of the water background also contained in the data as a reference (see ref 54 for details). Because the water background is the same in that light-induced difference spectrum, it did not have to be measured independently and a stationary cuvette with only ≈1 μL of sample volume was sufficient for these experiments. For the unfolding difference spectrum, however, L78AHA was resuspended in 6 M guanidinium chloride, deuterated buffer solution at pH 8.5 to ensure complete unfolding. The final spectrum is a double-difference spectrum, i.e., the spectrum in the unfolded state minus that of the corresponding buffer, subtracted from the folded state minus corresponding buffer. To subtract out the buffer contribution, a syringe pump sample

DOI: 10.1021/acs.jpca.7b09675
J. Phys. Chem. A 2017, 121, 9435−9445

delivery system together with a flow cell was used to exchange sample,[20,22] requiring much larger sample volumes of ≈100 μL.

II.C. Molecular Dynamics Simulations. Recently, Buchenberg et al. performed a detailed molecular dynamics (MD) study of the structural changes of photoswitchable PDZ2 upon *cis*−*trans* photoisomerization.[31] Following this work, we carried out MD simulations of photoswitchable PDZ2 including AHA labels as present in the experiment. To minimize computational time, three labels were considered per simulated system (in contrast to experiment, where only one amino acid was replaced per sample). Selecting label positions such that no AHA group interacts with another one, the first system contained AHA labels at sequence positions 16, 66, and 78, and the second system at positions 20, 48, and 80. Both systems were simulated in the *cis* and *trans* configuration of PDZ2, using Gromacs 4.6.7 with a hybrid GPU-CPU acceleration scheme.[55]

In all simulations, the protein was placed in aqueous solution including 150 mM NaCl. The side chains of all four histidine residues (33, 54, 72, and 87) in all initial structures were chosen to be ϵ-protonated. The protein was described using the Amber99SB*ILDN force field,[56−58] water molecules by the TIP3P model,[59] and ions with the model of ref 60. The parametrization procedure using Antechamber[61] and Gaussian09[62] and the resulting force field parameters of the AHA labels are described in the Supporting Information. All bonds involving hydrogen atoms were constrained,[63] allowing for a 2 fs time step. Electrostatic interactions were calculated using PME.[64] The minimum cutoff distance for electrostatic and van der Waals interactions was set to 1.2 nm. To couple the system to a heat bath, we used the velocity-rescale algorithm[65] and for pressure coupling the Berendsen algorithm.[66] After energy minimization, the systems were simulated for 100 ns at a pressure of 1 bar and a temperature of 300 K.

Data evaluation was carried out with Gromacs tools.[55] To determine intraprotein polar contacts, we used *g_mindist* to calculate the minimal distances between the $-N_3$ atoms of the AHA residue and all polar atoms found within a 1 nm radius of the azido group (with respect to the starting structure). Contact distributions were then obtained by histogramming the MD data with 0.01 nm binning width. We defined a contact to be formed if the minimal distance between a azido group nitrogen atom and a protein nitrogen or oxygen atom is shorter than 0.45 nm.[67] In a similar way, we analyzed contacts between AHA and water as azido group/water oxygen atom distances with a cutoff of 0.45 nm, as well.

II.D. Calculation of Vibrational Spectra. We used the empirical model of Cho and co-workers[28] to estimate vibrational shifts $\delta\omega$ caused by changes in the electrostatic environment of the AHA labels. By calculating the electric field $\mathbf{E}_j(t)$ at the nitrogen atoms (j = 1, 2, 3) of the azido group for each MD snapshot at time t, we obtain the spectral shift

$$\delta\omega(t) = \sum_j \mathbf{a}_j \mathbf{E}_j(t) \tag{1}$$

with coefficients $\mathbf{a}_j$ given in ref 28. Electric fields were computed via a reaction field approach[55,68] using a cutoff radius r_c = 2.3 nm. From the frequency trajectory $\delta\omega(t)$ with a time step of 15 ps, the distribution of the vibrational shifts was obtained via a histogram using 50 bins.

As an alternative approach, quantum-mechanical calculations of vibrational spectra of protein side-chain conformers were performed using Gaussian09,[62] following Wolf et al.[69] From the MD simulation, we first determined the snapshot where the $N^{(2)}$-atom of the azido group is closest to one of the two sulfur atoms of the photoswitch sulfonate groups; we chose one structure each for the *trans* and the *cis* configuration, and for I20AHA and L78AHA. Using these structures, we constructed a minimal vacuum model of the AHA label ($-N_3$ group and $C\beta$ atom) and the sulfonate group (including the attached carbon atom), where both carbon atoms were saturated with hydrogen atoms. Keeping the positions of both carbon atoms as well as the distance between $N^{(2)}$-atom and sulfur atoms fixed, the resulting system was initially minimized in energy at the HF/6-31+G* level, followed by a density functional theory-based minimization using B3LYP[70,71] and the 6-31+G* basis set. At the same theoretical level, harmonic frequencies and band intensities were determined by diagonalization of the Hessian matrix. Vibrational frequencies were corrected by the asymmetry factors given in ref 72. Atomic charges were calculated via Mulliken population analysis.[73]

III. RESULTS AND DISCUSSION

III.A. Folding Stability. We have determined the folding stability of the various mutants by CD spectroscopy; the data are shown in Figure S2 (Supporting Information), and the results are summarized in Table 1. The CD measurements have been performed under the same conditions as the 2D IR experiments, i.e., buffered in D_2O solution. Tentatively, the AHA mutations destabilize the protein, if at all, only a little bit relative to the photoswichtable PDZ2 domain without AHA, as judged from the midpoint temperatures T_m. The *cis*-state is, overall speaking, somewhat less stable than the *trans*-state (the latter is opposite to what we have reported in ref 30, where the CD has, however, been measured in H_2O and without NaCl). Furthermore, the folding transition is less cooperative in the *cis*-state with a larger width of the folding transitions ΔT. Nevertheless, all mutants considered here are folded to ≳98% at 10 °C in both their *cis*- and their *trans*-states. We performed the 2D IR experiments at that temperature to ensure that any difference signal induced by photoswitching is not obscured by partial unfolding.

Table 1. Unfolding Midpoint Temperatures T_m and Width of the Folding Transitions ΔT of the Different Mutations of the Photoswitchable PDZ2 Domain in Its Two States, As Obtained from CD Spectroscopy[a]

	trans		*cis*	
mutant	T_m/°C	ΔT/°C	T_m/°C	ΔT/°C
no AHA	49.5 ± 1	3.7 ± 1	46.0 ± 2	6.3 ± 2
N16AHA	48.0 ± 1	5.4 ± 1	45.5 ± 2	8.0 ± 2
I20AHA	43.0 ± 1	5.2 ± 1	36.5 ± 4	7.9 ± 2
S48AHA	49.5 ± 1	3.9 ± 1	51.0 ± 1	3.8 ± 1
L66AHA	44.5 ± 1	4.2 ± 1	43.5 ± 1	5.9 ± 1
L78AHA	43.5 ± 1	4.4 ± 1	38.5 ± 2	7.4 ± 2
N80AHA	47.0 ± 1	4.5 ± 1	49.0 ± 3	6.0 ± 3

[a] To that end, the data have been fit to a function $1/(1 + \exp((T - T_m)/\Delta T))$ after subtraction of the background and normalisation.

III.B. Unfolding 2D IR Difference Spectra. As a reference experiment, Figure 2 compares the 2D IR spectra of L78AHA in the folded and the unfolded state of the protein. The measurements were performed in the dark, i.e., with the photoswitch in its *trans* configuration. We chose to induce unfolding by a denaturant (6 M guanidinium chloride), rather than by raising the temperature, as the latter causes a dramatic

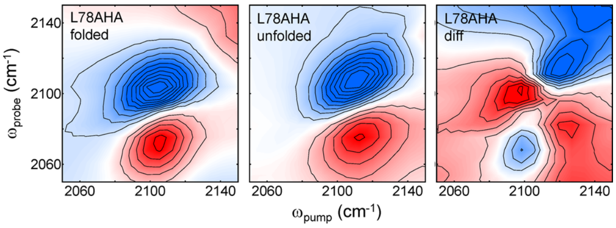


Figure 2. Purely absorptive 2D IR spectra of L78AHA in the folded (left panel) and the unfolded state (middle panel), the latter induced by adding 6 M guanidinium chloride. The right panel shows the difference (unfolded minus folded) of both spectra. To that end, the spectra of the folded and unfolded protein have been scaled to the bleach signal, because they have been measured at slightly different concentrations and have subsequently been subtracted to reveal the difference spectrum. For the plotting, the difference spectrum has been multiplied by a factor 2 so that all spectra share the same number of contour lines. Blue colors depict negative signals (i.e., bleach and stimulated emission in the purely absorptive 2D IR spectra) and red colors positive signals (excited state absorption).

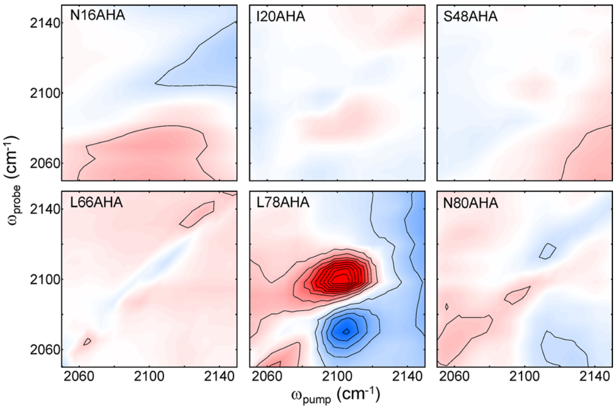


Figure 3. Purely absorptive 2D IR difference spectra (*cis* minus *trans*) of all mutants considered in this study, induced by photoswitching the azobenzene moiety from *trans* to *cis* with the help of a cw-laser diode. All signals were normalized by concentration (which varied between 0.6 and 1.1 mM) and to the peak signal of L78AHA. Blue colors depict negative signals and red colors positive signals.

change in the water response and furthermore tends to induce aggregation of the protein and hence scattering in the 2D IR signal. In each case, a 2D IR spectrum of the corresponding buffer has been measured as well under identical conditions and has been subtracted, as shown in refs 4, 20, and 22. In both states of the protein, the 2D IR spectra show the usual 0−1 peak depicted in blue together with the 1−2 peak depicted in red (i.e., with opposite sign), which is shifted along the probe-frequency axis due to the anharmonicity of the AHA vibration. By the tilt of the 2D IR lineshapes, a modest amount of inhomogeneity is detected, which does not differ very much in the two states of the protein.

The most prominent change upon unfolding is a blue shift of the AHA label by ≈7 cm^{-1}, which is a bit smaller than for previous observations.[17] That is, upon unfolding, the AHA label becomes fully solvated and hence the number of hydrogen bonds to water molecules increases, as well as their flexibility allowing for hydrogen bonding at a more optimal angle. Cho and co-workers have shown in ref 18 that both effects cause a blue shift of the vibrational transitions. In turn, the frequency shift also reveals that the AHA label of this particular mutant is solvent-exposed to only a minor extent in the folded state of the protein. Position L78AHA is situated inside the binding pocket (Figure 1), where the access of the solvent is limited, possibly shielded by the azobenzene photoswitch (see below).

III.C. Photoswitching 2D IR Difference Spectra. With that information in mind, we turn to the 2D IR difference spectroscopy induced by photoswitching the azobenzene moiety from *trans* to *cis* with the help of a cw-laser diode. Figure 3 shows the spectra of all mutants that have been investigated. Surprisingly, only one of the considered mutants (L78AHA) reveals an evaluable signal, while all others show no clear signal apart from some modifications of the water background and/or small remaining scattering (the latter appears as spurious signals along the diagonal; see, e.g., the 2D IR difference spectra of L66AHA or N80AHA). These two effects currently limit our sensitivity, not the signal-to-noise ratio *per se*. It should, however, be stressed that we know beyond any doubt that the AHA label is present also in those mutants, for which no difference signal could be detected. For example, we see its absorption band in the individual dark-adapted (*trans*) 2D IR spectra, i.e., before taking the difference with the corresponding *cis* 2D IR spectra (however, sitting on a large water background; see Figure S3, Supporting Information). Also mass spectrometry (Figure S1, Supporting Information) confirms the existence of an azido group in all mutants.

Nonetheless, the response of L78AHA, and the comparison to the unfolding difference spectrum from Figure 2, is quite revealing. The 2D IR difference spectrum induced by photoswitching (Figure 3) is quite comparable to the 2D IR spectrum of the folded state (Figure 2, left panel), where the frequency position and 2D line shape is concerned, but the sign (encoded by the colors) is inverted and the intensity is about a factor 5 smaller. Furthermore, the photoswitch induced difference spectrum is very different from the unfolding induced difference spectrum (Figure 2, right panel), the latter of which results from the frequency shift of the AHA transition. That is, the effect of photoswitching on the AHA label is mostly a reduction of the vibrational transition dipole in the *cis*-state without affecting the vibrational frequency very much. This in turn also evidences that the difference spectrum induced by photoswitching is not the result of the slightly reduced stability of the protein with the photoswitch in the *cis* state, which according to CD spectroscopy (Supporting Information, Figure S2) might cause ≈2% of unfolding.

To the best of our knowledge, this effect has not been described so far. That is, although the frequency of the AHA label is considered to be a measure of the amount of solvation,[16−22] the intensity stays essentially the same, which is indeed what is observed when the protein unfolds (Figure 2). It is conceivable that when the label enters a more heterogeneous environment, the absorption becomes wider at the expense of the peak intensity such that the integrated intensity, and hence the transition dipole, stays the same. We tentatively exclude that effect here, because it would cause wings of opposite sign on both sides of the peak in the difference spectrum that are not observed, even when considering the present signal-to-noise level. We also carefully checked the possibility that the loss of intensity reflects a loss of AHA label due to a reduction of the azido group, for example. To that end, we first measured mass spectra before and after laser illumination (Figure S4, Supporting Information),

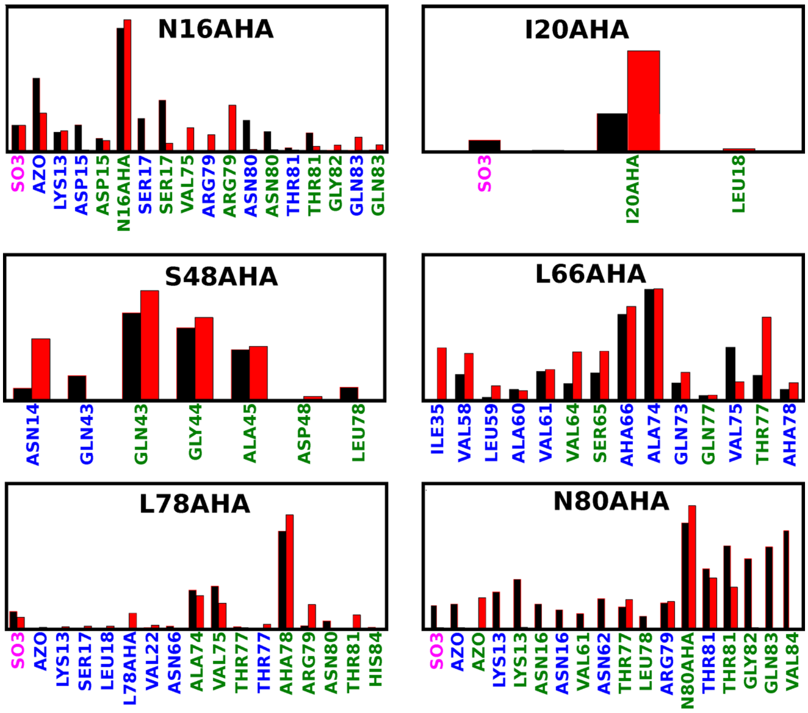


Figure 4. Histograms of polar contacts between considered AHA labels and PDZ2 protein in *cis* (black) and *trans* (red) states. Green and blue labels denote backbone and side-chain contacts, respectively, and magenta labels indicate contacts with the sulfonate groups of the photoswitch.

showing that no chemical modification is occurring, such as the loss of a N_2 molecule. Second, we measured FTIR difference spectra switching forth and back, evidencing that the transition is indeed reversible and that the AHA band regains its intensity upon *cis*−*trans* back-switching (Figure S5, Supporting Information).

III.D. Spectral Simulations. To get insights into the molecular mechanism giving rise to this spectroscopic response, we performed spectral simulations based on all-atom MD simulations. Recent quantum-chemical calculations of Cho and co-workers[18] have shown that the spectroscopic signatures of the azido stretch mode of AHA mainly reflect the local electrostatic environment of the azido group. To result in an observable IR difference spectrum of PDZ2, this electrostatic field needs to differ in the *trans* and the *cis* states. Because the main contributions to the local electrostatic environment arise from contacts with nearby polar residues and water molecules, we first employ MD simulations to study possible changes of these contacts caused by the *trans*−*cis* isomerization.

Let us begin with the contacts between polar protein residues and the azido groups. Figure 4 reveals that the number of such intraprotein contacts and their changes upon *trans*−*cis* photoisomerization appear quite complex. Roughly speaking, we find three labels (N16AHA, L66AHA, N80AHA) with a large number, two (S48AHA, L78AHA) with a medium number, and one (I20AHA) with a small number of intraprotein contacts. Notably, we see that labels N16AHA, I20AHA, and N80AHA show significant contact changes between *trans* and *cis* states, whereas the label L78AHA shows only minor variations. Figure 5 displays the distribution of water contacts during the MD simulation. Three AHA labels (N16AHA, S48AHA, N80AHA) are found to be strongly hydrated, two (I20AHA, L66AHA) are hardly hydrated, and label L78AHA, for which a difference signal could be observed experimentally, may be characterized as moderately hydrated. Interestingly, only the strongly hydrated labels show significant differences in the number of water contacts between *trans* and *cis* states. In particular, we find that the decrease in the number of water contacts of N16AHA and N80AHA upon *trans*−*cis* photoisomerization is compensated by an increase of protein contacts.

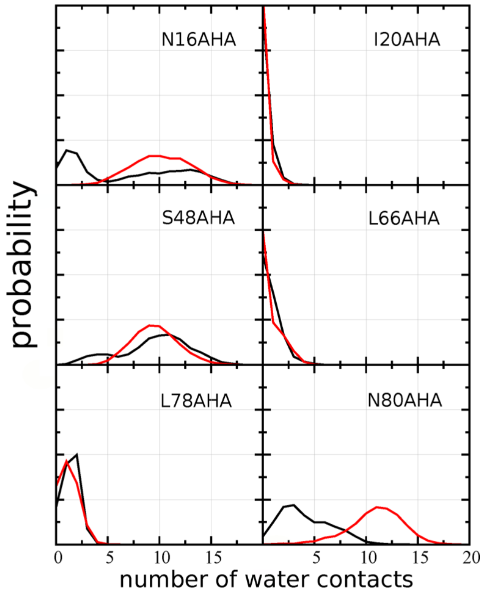


Figure 5. Distribution of the number of water contacts of the AHA labels, with the protein being in *trans* (red) and *cis* (black) state.

As the examination of the changes of contacts does not provide a clear explanation of the experimental findings, we next study to what extent the contact changes discussed above are reflected in spectral changes of the azido stretch mode of the corresponding AHA label. To this end, Figure 6 displays the distribution of vibrational frequency shifts $\delta\omega(t)$ upon *trans*−*cis* photoisomerization, as obtained from the electrostatic model in eq 1. In the case of the strongly hydrated labels (N16AHA, S48AHA, N80AHA), the calculations predict relatively broad ($\sim\pm10\ \mathrm{cm}^{-1}$) and red-shifted (by $\sim10\ \mathrm{cm}^{-1}$) frequency distributions. The weakly to hardly hydrated labels (I20AHA, L66AHA, L78AHA), however, exhibit a smaller distribution width ($\sim\pm5\ \mathrm{cm}^{-1}$) and also smaller frequency shifts. This different width of the frequency distributions of hydrated vs not hydrated labels is in nice agreement with the experimental results of Taskent-Sezgin et al.,[17] who found line widths of $\sim\pm5$ and $10\ \mathrm{cm}^{-1}$ for AHA labels in the folded and unfolded (i.e., water exposed) state of a protein, respectively. However, we note that the overall red shift predicted by the model is in variance with the common expectation that solvation rather causes a blue shift (see, e.g., Figure 2 as well as refs 1, 17, and 20−22). As discussed in ref 28 [Figure 6], this effect is caused by our neglect of the polarizability of the azido group.

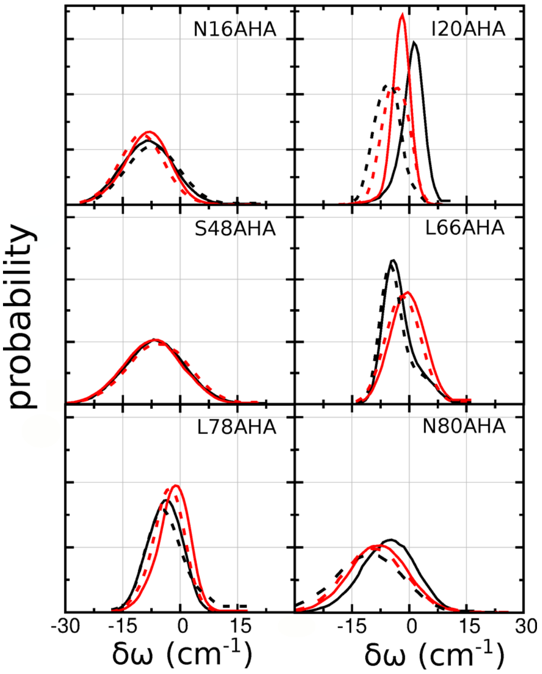


Figure 6. Vibrational frequency shifts (relative to the gas-phase vibration) of the azido stretch mode of the AHA labels in the *trans* (red) and *cis* (black) states. Results are obtained from the electrostatic model (eq 1) with (full lines) and without (dashed lines) inclusion of the sulfonate groups of the photoswitch.

Upon *trans*−*cis* photoisomerization, the strongly hydrated labels show no or only minor spectral changes. This is clearly expected for S48AHA, where both protein and water contacts remain more or less unchanged, but comes a bit as a surprise for N16AHA and N80AHA, where both protein and water contacts change significantly. We conclude that the highly mobile water molecules around these labels may effectively screen or counter the electrostatic interactions of the protein, which hampers a clear spectroscopic response. The frequency distribution of L78AHA hardly changes between *trans* and *cis* states, which is in line with the absence of major protein or water contact changes for this label. However, the spectral simulations would predict a significant blue shift of I20AHA and L66AHA when changing from *trans* to *cis*, resulting from the changes of protein contacts, which, however, is not observed experimentally.

To sum up the results up to this point, our MD simulations combined with the electrostatic model by Cho and co-workers[28] would predict spectral changes for I20AHA and L66AHA, whereas in experiment L78AHA is the only label with an observable spectral response upon *trans*−*cis* photoisomerization. Of course, one may question the accuracy of the structural prediction of the employed MD force field (neglecting, e.g., the polarizability of the azido group) or the assumptions underlying the electrostatic model (neglecting, e.g., the dependence on dispersive interactions or the fact that the model of Cho and co-workers[28] has been parametrized on the basis of QM calculations with water clusters around an azido group only, whereas intraprotein contact to polar residue play a significant role here as well). Moreover, the electrostatic model would describe frequency shifts only, which is not what is observed experimentally. That is, the experimental 2D IR difference spectrum upon *trans*−*cis* photoisomerization is dominated by a change in the intensity of the vibration of L78AHA, and not its frequency.

Being a 1,3-dipole, the azido group is strongly polarizable, and it is, for example, well-known that azido groups in different molecules have strongly varying transition dipoles.[5] Hence, it is conceivable that even weak electronic interactions with other parts of the protein, involving, for example, a charge transfer similar to the situation in a hydrogen bond or at the onset of a nucleophilic attack, may strongly affect the electronic density of the azido group and thereby change the transition dipole of its vibration. The prime candidates for such an electronic interaction are the two sulfonate groups of the photoswitch, which have been introduced to increase the solubility of the azobenzene moiety.[49] When electrostatic calculations are performed in the absence of these sulfonate groups (dashed lines in Figure 6), the frequency distributions of I20AHA and L78AHA reveal a significant red shift, emphasizing the importance of the sulfonate groups. Moreover, the red shifts are clearly different for *trans*- and *cis*-states. The strongly hydrated labels N16AHA and N80AHA, however, show only minor spectral changes when sulfonate groups are excluded in the electrostatic calculations. Although these labels form contacts with the sulfonate groups (Figure 4), the strong electrostatic screening due to solvent water appears to prevent clear spectral changes.

In our MD simulations, such contacts do not appear often (cf. Figure 4), owing to the missing stabilizing electronic interaction in the force field description. To investigate if such direct contacts between azido and sulfonate groups can lead to a charge transfer and the experimentally observed absorbance changes, we extracted the structure with the closest distance between I20AHA/L78AHA and a neighboring sulfonate group from MD simulations of each *cis* and *trans* states. Showing the energy-minimized structures of these states, parts a and b of Figure 7 reveal that L78AHA indeed forms a close contact with the sulfonate group, which mediates an electronic interaction. Intriguingly, this contact only exists in the *cis* configuration of the photoswitch, but not in the *trans* configuration. Likewise, an energy-minimized structure of I20AHA (not shown) exhibits a close contact associated with an electronic interaction in the

DOI: 10.1021/acs.jpca.7b09675
J. Phys. Chem. A 2017, 121, 9435−9445

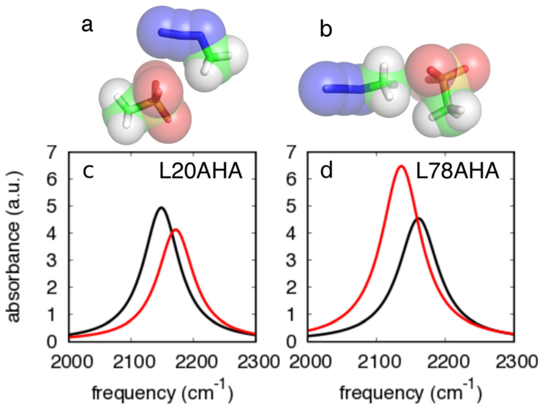


Figure 7. Quantum-chemical calculations for a minimal model of the AHA label ($-N_3$ group and Cβ atom) and the sulfonate group of the photoswitch, indicating the van der Waals radii of the atoms by colored spheres. Panels a and b show energy-minimized structures of L78AHA in *cis* with contact and *trans* without contact, respectively, of the azido and the sulfonate groups. I20AHA exhibited similar structures, with a contact in *trans* and without contact in *cis* (not shown). Using these structures, vibrational spectra of the azido stretch mode were calculated for (c) I20AHA and (d) L78AHA. Black and red lines correspond to the *cis*- and *trans*-state of the photoswitch, respectively.

trans configuration, which does not exist in the *cis* configuration.

Using these structures, we calculated harmonic frequencies and band intensities of the azido stretch mode (Materials and Methods). Spectra c and d in Figure 7 show the resulting vibrational spectra of the azido stretch mode of I20AHA and L78AHA in the *cis* and *trans* states of the photoswitch. In either case, a close contact between the azido and the sulfonate groups results in a decrease of the band intensity of the AHA vibration due to a charge transfer to the azido group that reduces the transition dipole. As can be seen in Table 2, the charge on all nitrogen atoms changes by up to ±0.23 *e*, indicating that the charges on the azido group are indeed very mobile due to electronic effects. A polarizable force field is the least that would be needed to describe that effect self-consistently in a MD simulation. The effect on intensity is stronger for L78AHA than for I20AHA, in qualitative agreement with experiment. It should, however, be mentioned that in addition to the intensity change, the quantum chemistry calculations would also predict a frequency shift of ~25 cm^{-1}, which is not observed in experiment. Possibly, the missing protein environment and solvation in the quantum chemistry calculations might be responsible for that effect. The results nevertheless suggest that weak electronic interactions may indeed change the intensity of the AHA vibration, enabling a new application for that vibrational label.

Table 2. Mulliken Charges[73] (in Units of *e*) of the AHA Azido Nitrogen Atoms in Different Quantum Chemical Vacuum Models[a]

	I20AHA		L78AHA	
	cis	*trans*	*cis*	*trans*
nitrogen	no contact	contact	contact	no contact
$N^{(3)}$	−0.45	−0.33	−0.52	−0.50
$N^{(2)}$	1.02	1.21	1.08	0.89
$N^{(1)}$	−0.71	−0.94	−0.67	−0.54

[a]Cont/no cont refers to the presence/absence of a contact with the neighboring sulfonate group. $N^{(3)}$: nitrogen attached to AHA side chain. $N^{(2)}$: internal nitrogen atom. $N^{(1)}$: terminal nitrogen atom.

IV. CONCLUSION

The AHA vibration is an attractive label to study protein structure and dynamics with the help of IR and 2D IR spectroscopy, because it can be incorporated almost anywhere in a protein by a methionine auxotrophic protein expression strategy,[26,27] because its transition dipole is reasonably large for 2D IR measurements in the sub-millimolar regime,[4,22] and because it is a sensitive probe of its solvation environment.[1,17−22,28] However, to demonstrate the last aspect, rather dramatic changes have been applied to the label so far. For example, a protein has been unfolded, which fully exposes an AHA label that normally is situated in the hydrophobic core of the protein to water,[17] or a peptide ligand with an AHA label has been dissociated from a protein,[20,22] which has about the same impact on the label as unfolding. Employing a combined experimental−computational approach, we set out in the present work to explore the capability of the AHA label to also sense much smaller changes in the structure and solvation of a midsized protein. That is, rather than unfolding of a protein, we change the structure of a PDZ2 domain in a very modest way (≲1 Å) with the help of a photoswitch that is covalently linked to it, thereby mimicking the conformational transition upon ligand binding.[29,31,32]

To that end, we first had to develop a protocol for the post-translational synthesis of the cross-linked PDZ2 domain containing AHA labels. That is, both the cysteins needed for the linking of the azobenzene photoswitch as well as the azido groups of AHA are chemically very reactive groups. For AHA, this reactivity has been used successfully for click chemistry reactions;[74] however, any modification of this unnatural amino acid had to be avoided in our case. Mainly, the reduction of disulfide bridges had to be optimized, such that AHA would not be reduced at the same time into a primary amine.

We have explored six mutations, distributing the AHA label at various positions of the protein that differ significantly in their properties (Figure 1), e.g., inside the hydrophobic core versus surface exposed, or in the flexible loops versus in more rigid secondary structure motives. We find that because of the fact that the label is very small and of medium polarity, it can replace both polar and apolar amino acids without affecting the stability of the protein too much (Table 1). In that sense, it is indeed a versatile label.

Somewhat surprisingly, however, we observe an evaluable difference 2D IR signal upon photoswitching only for one mutation, L78AHA (we nevertheless chose to show all results in Figure 3, including the negative ones, because the purpose of this survey has been to learn what can be sensed with the AHA label, and what its limitations are). The very distinct difference signal of only the one label L78AHA comes as a surprise, because I20AHA and N80AHA are equally close to perturbation introduced by the photoswitch. Naturally, one would assume that the size of the effect correlates with the distance to the perturbation; yet, with our current experimental sensitivity we cannot detect any response for I20AHA and N80AHA. It also comes as a surprise, because the MD simulations would in fact predict significant changes of local contacts to both water molecules and protein side chains, even

DOI: 10.1021/acs.jpca.7b09675
J. Phys. Chem. A 2017, 121, 9435−9445

for labels that are quite far away from the photoswitch, such as N16AHA or L66AHA (Figures 4 and 5).

The distinct response of L78AHA, in turn, suggests that it results from a very specific interaction that should be robustly reproduced by a MD simulation, even given the unavoidable imperfectness of any MD force field that might describe some structural details in not quite the correct way. We therefore first employed classical MD simulations, calculating local contacts of the various AHA labels with its surrounding (Figures 4 and 5) as well as electric field induced frequency shifts along the line of a model put forward by Cho and co-workers (Figure 6).[28] The simulation revealed that the strongly hydrated labels show no or only minor spectral changes, even if both protein and water contacts change significantly (as is the case for N16AHA and N80AHA). This suggests that this is the result of a mutual cancellation of the contributions from water and the protein. That is, the highly mobile water molecules around these labels effectively screen or counter the electrostatic interactions of the protein.

However, the calculations revealed nothing with respect to which L78AHA would stick out. In the contrary, they predict frequency shifts for I20AHA and L66AHA that are of the same order as in the unfolding experiment of Figure 2 or what has been observed in ref 22 for ligand unbinding and thus should be measurable with our current sensitivity. One possible explanation for that discrepancy might be the fact that the electrostatic model of Cho and co-workers[28] neglects polarizability and has been parametrized on the basis of quantum chemical calculations focusing on hydrogen bonding in water clusters around a solvated azido group only, whereas L66AHA and I20AHA are basically not solvated (Figure 5) and the electric field in eq 1 originates mostly from intrapeptide contact with polar side groups.

To explain the distinct response of L78AHA, which actually reflects a change of intensity of the AHA vibration rather than its frequency, we speculate here that it originates from an electronic interaction between the AHA label and the sulfonate groups of the azobenzene moiety. As a proof of principle, preliminary gas-phase quantum chemistry calculations, using structures derived from the classical MD simulation, give evidence that this might indeed be the case (Figure 7). We must concede that we did not do these calculations on a sample of possible contact structures, that the agreement with experiment is rather modest, and that a polarizable force field or QM/MM simulations would be needed to verify that mechanism. As the interaction of the azido group with the sulfonate group is related to a weak chemical bond, the effect might be amplified in a QM/MM simulation, thus increasing the stability and occurrence of contact structures. The same seems to happen also upon hydrogen bonding to water, as evidenced by the observation of Cho and co-workers[18,28] that QM/MM simulations are necessary to correctly descibe the water structure around the azido group, which is a prerequisite for predicting a blue shift upon solvation.

Despite the preliminary character, the current results nevertheless show that the intensity of the AHA vibration carries information about its surrounding that is complementary to the frequency position. This effect has not been described in the literature to the best of our knowledge and needs to be taken into account in the analysis of difference spectra from IR labels that can undergo electronic interactions with neighboring amino acids, such as azido and cyano groups. In a broader sense, it reminds one that in difference spectroscopy, band disappearances do not necessarily need to be an effect of a vanishing state population but can as well stem from effects altering the transition dipole moment of the reporting label. In conclusion, we have found a new and unexpected mode of spectral response of an AHA label incorporated into a protein. This response stems from an electronic instead of an electrostatic interaction of the AHA azido group with the protein/water surrounding, and opens up new applications for the usage of artificial probes in biospectroscopy.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.7b09675.

> Details of the force field parametrization of the AHA group, as well as mass spectra determining the purity of the samples, temperature-dependent CD spectra determining their stability, 2D IR spectra of the dark-adapted (*trans*) state, and FTIR difference spectra (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*P. Hamm. E-mail: peter.hamm@chem.uzh.ch.
*G. Stock. E-mail: stock@physik.uni-freiburg.de.

ORCID

Peter Hamm: 0000-0003-1106-6032
Gerhard Stock: 0000-0002-3302-3044

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Ben Schuler and his group for their continuous and tremendous help with the protein chemistry, Brankica Janković for helpful discussion regarding the sample preparation, Steven Waldauer for his idea to use magnetic beads for a faster protein purification, and the Functional Genomics Center Zurich, especially Serge Chesnov, for help with the mass spectrometry. The work has been supported in part by a European Research Council (ERC) Advanced Investigator Grant (DYNALLO), by the Swiss National Science Foundation (SNF) through the NCCR MUST and Grant 200021_165789/1, as well as by the Deutsche Forschungsgemeinschaft through Grant STO 247/10-1. Computational resources were provided by the bwForCluster BinAC (RV bw16I016) and the Black Forest Grid Initiative.

■ REFERENCES

(1) Waegele, M. M.; Culik, R. M.; Gai, F. Site-Specific Spectroscopic Reporters of the Local Electric Field, Hydration, Structure, and Dynamics of Biomolecules. *J. Phys. Chem. Lett.* **2011**, *2*, 2598−2609.

(2) Adhikary, R.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. IR Probes of Protein Microenvironments: Utility and Potential for Perturbation. *ChemPhysChem* **2014**, *15*, 849−853.

(3) Kim, H.; Cho, M. Infrared Probes for Studying the Structure and Dynamics of Biomolecules. *Chem. Rev.* **2013**, *113*, 5817−5847.

(4) Koziol, K. L.; Johnson, P. J. M.; Stucki-Buchli, B.; Waldauer, S. A.; Hamm, P. Fast Infrared Spectroscopy of Protein Dnamics: Advancing Sensitivity and Selectivity. *Curr. Opin. Struct. Biol.* **2015**, *34*, 1−6.

(5) Ma, J.; Pazos, I. M.; Zhang, W.; Culik, R. M.; Gai, F. Site-Specific Infrared Probes of Proteins. *Annu. Rev. Phys. Chem.* **2015**, *66*, 357−377.

(6) Koziński, M.; Garrett-Roe, S.; Hamm, P. 2D-IR Spectroscopy of the Sulfhydryl Band of Cysteines in the Hydrophobic Core of Proteins. *J. Phys. Chem. B* **2008**, *112*, 7645−7650.

(7) Naraharisetty, S. R. G.; Kasyanenko, V. M.; Zimmermann, J.; Thielges, M.; Romesberg, F. E.; Rubtsov, I. V. C-D Modes of Deuterated Side Chain of Leucine as Structural Reporters via Dual-frequency Two-dimensional Infrared Spectroscopy. *J. Phys. Chem. B* **2009**, *113*, 4940−4946.

(8) Schade, M.; Moretto, A.; Crisma, M.; Toniolo, C.; Hamm, P. Vibrational Energy Transport in Peptide Helices after Excitation of C-D Modes in Leu-D10. *J. Phys. Chem. B* **2009**, *113*, 13393−13397.

(9) Zimmermann, J.; Thielges, M. C.; Yu, W.; Dawson, P. E.; Romesberg, F. E. Carbon-Deuterium Bonds as Site-Specific and Nonperturbative Probes for Time-Resolved Studies of Protein Dynamics and Folding. *J. Phys. Chem. Lett.* **2011**, *2*, 412−416.

(10) Anfinrud, P. A.; Han, C.; Hochstrasser, R. M. Direct Observations of Ligand Dynamics in Hemoglobin by Subpicosecond Infrared-Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 8387−8391.

(11) Lim, M.; Jackson, T. A.; Anfinrud, P. A. Mid-Infrared Vibrational Spectrum of CO after Photodissociation from Haem: Evidence for a Ligand Docking Site in the Haem Pocket of Haemoglobin and Myoglobin. *J. Chem. Phys.* **1995**, *102*, 4355−4366.

(12) Thielges, M. C.; Fayer, M. D. Protein Dynamics Studied with Ultrafast Two-Dimensional Infrared Vibrational Echo Spectroscopy. *Acc. Chem. Res.* **2012**, *45*, 1866−1874.

(13) Woys, A. M.; Mukherjee, S. S.; Skoff, D. R.; Moran, S. D.; Zanni, M. T. A Strongly Absorbing Class of Non-Natural Labels for Probing Protein Electrostatics and Solvation with FTIR and 2D IR Spectroscopies,. *J. Phys. Chem. B* **2013**, *117*, 5009−5018.

(14) Peran, I.; Oudenhoven, T.; Woys, A. M.; Watson, M. D.; Zhang, T. O.; Carrico, I.; Zanni, M. T.; Raleigh, D. P. General Strategy for the Bioorthogonal Incorporation of Strongly Absorbing, Solvation-Sensitive Infrared Probes into Proteins. *J. Phys. Chem. B* **2014**, *118*, 7946−7953.

(15) King, J. T.; Arthur, E. J.; Brooks, C. L.; Kubarych, K. J. Crowding Induced Collective Hydration of Biological Macromolecules Over Extended Distance. *J. Am. Chem. Soc.* **2014**, *136*, 188−194.

(16) Oh, K.-I.; Lee, J.-H.; Joo, C.; Han, H.; Cho, M. β-Azidoalanine as an IR Probe: Application to Amyloid Aβ(16−22) Aggregation. *J. Phys. Chem. B* **2008**, *112*, 10352−10357.

(17) Taskent-Sezgin, H.; Chung, J.; Banerjee, P. S.; Nagarajan, S.; Dyer, R. B.; Carrico, I.; Raleigh, D. P. Azidohomoalanine: A Conformationally Sensitive IR Probe of Protein Folding Protein Structure and Electrostatics. *Angew. Chem., Int. Ed.* **2010**, *49*, 7473−7475.

(18) Choi, J. H.; Raleigh, D.; Cho, M. Azido Homoalanine is a Useful Infrared Probe for Monitoring Local Electrostatistics and Side-Chain Solvation in Proteins. *J. Phys. Chem. Lett.* **2011**, *2*, 2158−2162.

(19) Thielges, M. C.; Axup, J. Y.; Wong, D.; Lee, H. S.; Chung, J. K.; Schultz, P. G.; Fayer, M. D. Two-Dimensional IR Spectroscopy of Protein Dynamics Using Two Vibrational Labels: A Site-Specific Genetically Encoded Unnatural Amino Acid and an Active Site Ligand,. *J. Phys. Chem. B* **2011**, *115*, 11294−11304.

(20) Bloem, R.; Koziol, K.; Waldauer, S.; Buchli, B.; Walser, R.; Samatanga, B.; Jelesarov, I.; Hamm, P. Ligand Binding Studied by 2D IR Spectroscopy Using the Azidohomoalanine Label. *J. Phys. Chem. B* **2012**, *116*, 13705−13712.

(21) Wolfshorndl, M. P.; Baskin, R.; Dhawan, I.; Londergan, C. H. Covalently Bound Azido Groups are Very Specific Water Sensors, Even in Hydrogen-Bonding Environments. *J. Phys. Chem. B* **2012**, *116*, 1172−1179.

(22) Johnson, P. J. M.; Koziol, K. L.; Hamm, P. Quantifying Biomolecular Recognition with Site-Specific 2D Infrared Probes. *J. Phys. Chem. Lett.* **2017**, *8*, 2280−2284.

(23) Bagchi, S.; Boxer, S. G.; Fayer, M. D. Ribonuclease S Dynamics Measured Using a Nitrile Label with 2D IR Vibrational Echo Spectroscopy. *J. Phys. Chem. B* **2012**, *116*, 4034−4042.

(24) Zimmermann, J.; Thielges, M. C.; Seo, Y. J.; Dawson, P. E.; Romesberg, F. E. Cyano Groups as Probes of Protein Microenvironments and Dynamics. *Angew. Chem., Int. Ed.* **2011**, *50*, 8333−8337.

(25) van Wilderen, L. J. G. W.; Kern-Michler, D.; Müller-Werkmeister, H. M.; Bredenbeck, J. Vibrational Dynamics and Solvatochromism of the Label SCN in Various Solvents and Hemoglobin by Time Dependent IR and 2D-IR Spectroscopy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 19643−19653.

(26) Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. Incorporation of Azides into Recombinant Proteins for Chemoselective Modification by the Staudinger Ligation. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 19.

(27) Simon, M.; Zangemeister-Wittke, U.; Plückthun, A. Facile Double-Functionalization of Designed Ankyrin Repeat Proteins Using Click and Thiol Chemistries. *Bioconjugate Chem.* **2012**, *23*, 279−286.

(28) Choi, J.-H.; Oh, K.-I.; Cho, M. Azido-Derivatized Compounds as IR Probes of Local Electrostatic Environment: Theoretical Studies. *J. Chem. Phys.* **2008**, *129*, 174512.

(29) Buchli, B.; Waldauer, S. A.; Walser, R.; Donten, M. L.; Pfister, R.; Blöchliger, N.; Steiner, S.; Caflisch, A.; Zerbe, O.; Hamm, P. Kinetic Response of a Photoperturbed Allosteric Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 11725−11730.

(30) Waldauer, S. A.; Stucki-Buchli, B.; Frey, L.; Hamm, P. Effect of Viscogens on the Kinetic Response of a Photoperturbed Allosteric Protein. *J. Chem. Phys.* **2014**, *141*, 22D514.

(31) Buchenberg, S.; Knecht, V.; Walser, R.; Hamm, P.; Stock, G. Long-Range Conformational Transition of a Photoswitchable Allosteric Protein: A Molecular Dynamics Simulation Study,. *J. Phys. Chem. B* **2014**, *118*, 13468−13476.

(32) Buchenberg, S.; Sittel, F.; Stock, G. Time-Resolved Observation of Protein Allosteric Communication. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E6804−E6811.

(33) Kozlov, G.; Gehring, K.; Ekiel, I. Solution Structure of the PDZ2 Domain from Human Phosphatase hPTP1E and its Interactions with C-Terminal Peptides from the Fas Receptors. *Biochemistry* **2000**, *39*, 2572−2580.

(34) Kozlov, G.; Banville, D.; Gehring, K.; Ekiel, I. Solution Structure of the PDZ2 Domain from Cytosolic Human Phosphatase hPTP1E Complexed with a Peptide Reveals Contribution of the Beta 2-Beta 3 Loop to PDZ Domain-Ligand Interactions. *J. Mol. Biol.* **2002**, *320*, 813−820.

(35) Zhang, J.; Sapienza, P. J.; Ke, H.; Chang, A.; Hengel, S. R.; Wang, H.; Phillips, G. N.; Lee, A. L. Crystallographic and Nuclear Magnetic Resonance Evaluation of the Impact of Peptide Binding to the Second PDZ Domain of Protein Tyrosine Phosphatase 1E. *Biochemistry* **2010**, *49*, 9280−9291.

(36) Fuentes, E. J.; Der, C. J.; Lee, A. L. Ligand-Dependent Dynamics and Intramolecular Signalling in a PDZ Domain. *J. Mol. Biol.* **2004**, *335*, 1105−1115.

(37) Fuentes, E. J.; Gilmore, S. A.; Mauldin, R. V.; Lee, A. L. Evaluation of Energetic and Dynamic Coupling Networks in a PDZ Domain Protein. *J. Mol. Biol.* **2006**, *364*, 337−351.

(38) Gianni, S.; Walma, T.; Arcovito, A.; Calosci, N.; Bellelli, A.; Engström, A.; Travaglini-Allocatelli, C.; Brunori, M.; Jemth, P.; Vuister, G. W. Demonstration of Long-Range Interactions in a PDZ Domain by NMR, Kinetics, and Protein Engineering. *Structure* **2006**, *14*, 1801−1809.

(39) Ota, N.; Agard, D. A. Intramolecular Signaling Pathways Revealed by Molecular Anisotropic Thermal Diffusion. *J. Mol. Biol.* **2005**, *351*, 345−354.

(40) De Los Rios, P.; Cecconi, F.; Pretre, A.; Dietler, G.; Michielin, O.; Piazza, F.; Juanico, B. Functional Dynamics of PDZ Binding Domains: a Normal-Mode Analysis. *Biophys. J.* **2005**, *89*, 14−21.

(41) Sharp, K.; Skinner, J. J. Pump-Probe Molecular Dynamics as a Tool for Studying Protein Motion and Long Range Coupling. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 347−361.

(42) Dhulesia, A.; Gsponer, J.; Vendruscolo, M. Mapping of Two Networks of Residues that Exhibit Structural and Dynamical Changes

upon Binding in a PDZ Domain Protein. *J. Am. Chem. Soc.* **2008**, *130*, 8931−8939.

(43) Kong, Y.; Karplus, M. Signaling Pathways of PDZ2 Domain: a Molecular Dynamics Interaction Correlation Analysis. *Proteins: Struct., Funct., Genet.* **2009**, *74*, 145−154.

(44) Gerek, Z. N.; Ozkan, S. B. Change in Allosteric Network Affects Binding Affinities of PDZ Domains: Analysis Through Perturbation Response Scanning. *PLoS Comput. Biol.* **2011**, *7*, e1002154.

(45) Cilia, E.; Vuister, G. W.; Lenaerts, T. Accurate Prediction of the Dynamical Changes within the Second PDZ Domain of PTP1e. *PLoS Comput. Biol.* **2012**, *8*, e1002794.

(46) Lockless, S. W.; Ranganathan, R. Evolutionarily Conserved Pathways of Energetic Connectivity in Protein Families. *Science* **1999**, *286*, 295−299.

(47) Suel, G. M.; Lockless, S. W.; Wall, M. A.; Ranganathan, R. Evolutionarily Conserved Networks of Residues Mediate Allosteric Communication in Proteins. *Nat. Struct. Biol.* **2003**, *10*, 59−69.

(48) Chi, C. N.; Elfström, L.; Shi, Y.; Snäll, T.; Engström, A.; Jemth, P. Reassessing a Sparse Energetic Network Within a Single Protein Domain. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 4679−4684.

(49) Burns, D. C.; Zhang, F.; Woolley, G. A. Synthesis of 3,3′-bis(Sulfonato)-4,4’-bis(Chloroacetamido)azobenzene and Cysteine Cross-Linking for Photo-Control of Protein Conformation and Activity. *Nat. Protoc.* **2007**, *2*, 251−258.

(50) Hamm, P.; Kaindl, R. A.; Stenger, J. Noise Suppression in Femtosecond Mid-Infrared Light Sources. *Opt. Lett.* **2000**, *25*, 1798−1800.

(51) Volkov, V.; Schanz, R.; Hamm, P. Active Phase Stabilization in Fourier-Transform Two-Dimensional Infrared Spectroscopy. *Opt. Lett.* **2005**, *30*, 2010−2012.

(52) Middleton, C. T.; Strasfeld, D. B.; Zanni, M. T. Polarization Shaping in the Mid-IR and Polarization-Based Balanced Heterodyne Detection with Application to 2D IR Spectroscopy, Opt. *Opt. Express* **2009**, *17*, 14526−14533.

(53) Bloem, R.; Garrett-Roe, S.; Strzalka, H.; Hamm, P.; Donaldson, P. Enhancing Signal Detection and Completely Eliminating Scattering Using Quasi-Phase-Cycling in 2D IR Experiments, Opt. *Opt. Express* **2010**, *18*, 27067−27078.

(54) Johnson, P. J. M.; Koziol, K. L.; Hamm, P. Intrinsic Phasing of Heterodyne-Detected Multidimensional Infrared Spectra, Opt. *Opt. Express* **2017**, *25*, 2928−2938.

(55) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; Van Der Spoel, D.; et al. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit,. *Bioinformatics* **2013**, *29*, 845−854.

(56) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712−725.

(57) Best, R. B.; Hummer, G. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004−9015.

(58) Lindorff-Larsen, K. L.; Piana, S.; Palmo, K.; et al. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.

(59) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(60) Joung, I. S.; Cheatham, T. E. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112*, 9020−9041.

(61) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.

(62) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*; Gaussian Inc.: Wallingford, CT, 2009.

(63) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.

(64) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An Nlog(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(65) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling Through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(66) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(67) Ernst, M.; Sittel, F.; Stock, G. Contact- and Distance-Based Principal Component Analysis of Protein Dynamics. *J. Chem. Phys.* **2015**, *143*, 244114.

(68) Steinhauser, O. Reaction Field Simulation of Water. *Mol. Phys.* **1982**, *45*, 335−348.

(69) Wolf, S.; Freier, E.; Cui, Q.; Gerwert, K. Infrared Spectral Marker Bands Characterizing a Transient Water Wire Inside a Hydrophobic Membrane Protein. *J. Chem. Phys.* **2014**, *141*, 22D524.

(70) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.

(71) Becke, A. D. Density-Functional Thermochemistry III: the Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(72) Johnson, R., III. NIST 101. Computational Chemistry Comparison and Benchmark Database. *CCCBDBD Computational Chemistry Comparison and Benchmark Database*; NIST, 1999.

(73) Mulliken, R. S. Electronic Population Analysis on LCAO−MO Molecular Wave Functions. I. *J. Chem. Phys.* **1955**, *23*, 1833−1840.

(74) Moses, J. E.; Moorhouse, A. D. The Growing Applications of Click Chemistry,. *Chem. Soc. Rev.* **2007**, *36*, 1249−1262.

Supplementary Material: 2D-IR Spectroscopy of an AHA Labelled Photoswitchable PDZ2 Domain

Brigitte Stucki-Buchli, Philip Johnson, Olga Bozovic, Claudio Zanobini, Klemens Koziol, Peter Hamm*
Department of Chemistry, University of Zurich, Zurich, Switzerland
and
Adnan Gulzar, Steffen Wolf, Sebastian Buchenberg, Gerhard Stock
Institute of Physics, Albert Ludwigs University, Freiburg, Germany

I. AHA MD PARAMETERS

AHA parameterization procedure

Force field parameters for the azobenzene photoswitch and the attached cysteine side chains as well as for AHA labels were obtained with the Antechamber package.[1] For the determination of atomic charges, the structures of the switch with attached cysteine side chains in *cis* and *trans* conformations as well as the AHA labels were optimized on B3LYP/6-31G* level using the GAUSSIAN g09 program suite.[2] Atomic charges of the different conformers were then computed as Mulliken charges from HF/6-31G* single point calculations. Point charges for MD calculations were then obtained from multiconformational restrained electrostatic potential (RESP)[3] calculations. The Cβ partial charges of covalently attached Cys residues were constrained to the value given in the Amber99sb*ILDN force field.

AHA parameters

[AHA]
[atoms]

N	N	-0.41570	1
H	H	0.27190	2
CA	CT	0.07730	3
HA	H1	0.07950	4
CB	CT	-0.11720	5
HB1	H1	0.05338	6
HB2	H1	0.05338	7
CG	CT	0.15260	8
HG1	H1	0.06012	9
HG2	H1	0.06012	10
N1	Nah	-0.44965	11
N2	Nbh	0.20880	12
N3	Nch	-0.06395	13
C	C	0.59730	14
O	O	-0.56790	15

```
[ bonds ]

        -C      N
        N       H
        N       CA
        CA      HA
        CA      CB
        CA      C
        CB      HB1
        CB      HB2
        CB      CG
        CG      HG1
        CG      HG2
        CG      N1
        N1      N2
        N2      N3
        C       O

[ impropers ]

      -C        CA        N         H
      CA        +N        C         O

[ atomtypes ]

  Nah    7     14.01    0.0000    A      3.25000e-01   7.11280e-01
  Nbh    7     14.01    0.0000    A      3.25000e-01   7.11280e-01
  Nch    7     14.01    0.0000    A      3 .25000e-01  7.11280e-01

[ bondtypes ]

 Nah    CT     1        0.14770    262590.0 ; AHA side chain GAFF
 Nah    Nbh    1        0.12160    717470.0 ; AHA side chain GAFF
 Nbh    Nch    1        0.11240    992300.0 ; AHA side chain GAFF

[ angletypes ]

  CT   CT   Nah   1    109.500   555.640 ; CB-CG1-N1 AHA GAFF
  CT   Nah  Nbh   1    116.200   574.380 ;CG1-N1-N2 AHA GAFF
  Nah  CT   H1    1    109.600   412.210 ; N1-CG1-HG12 AHA GAFF
  Nah  Nbh  Nch   1    180.000   566.100 ; N1-N2-N3 AHA GAFF

[ dihedraltypes ]

      CT   CT   CT   Nah  3   0.65084 1.95253 0.00000   -2.60338   0.00000   0.00000
      CT   CT   Nah  Nbh  3   0.00000 0.00000 0.00000   0.00000    0.00000   0.00000
      CT   Nah  Nbh  Nch  3   0.00000 0.00000 0.00000   0.00000    0.00000   0.00000
      Nah  CT   CT   H1   3   0.65084 1.95253 0.00000   -2.60338   0.00000   0.00000
      Nbh  Nah  CT   H1   3   0.00000 0.00000 0.00000   0.00000    0.00000   0.00000
```

II. SUPPLEMENTARY FIGURES

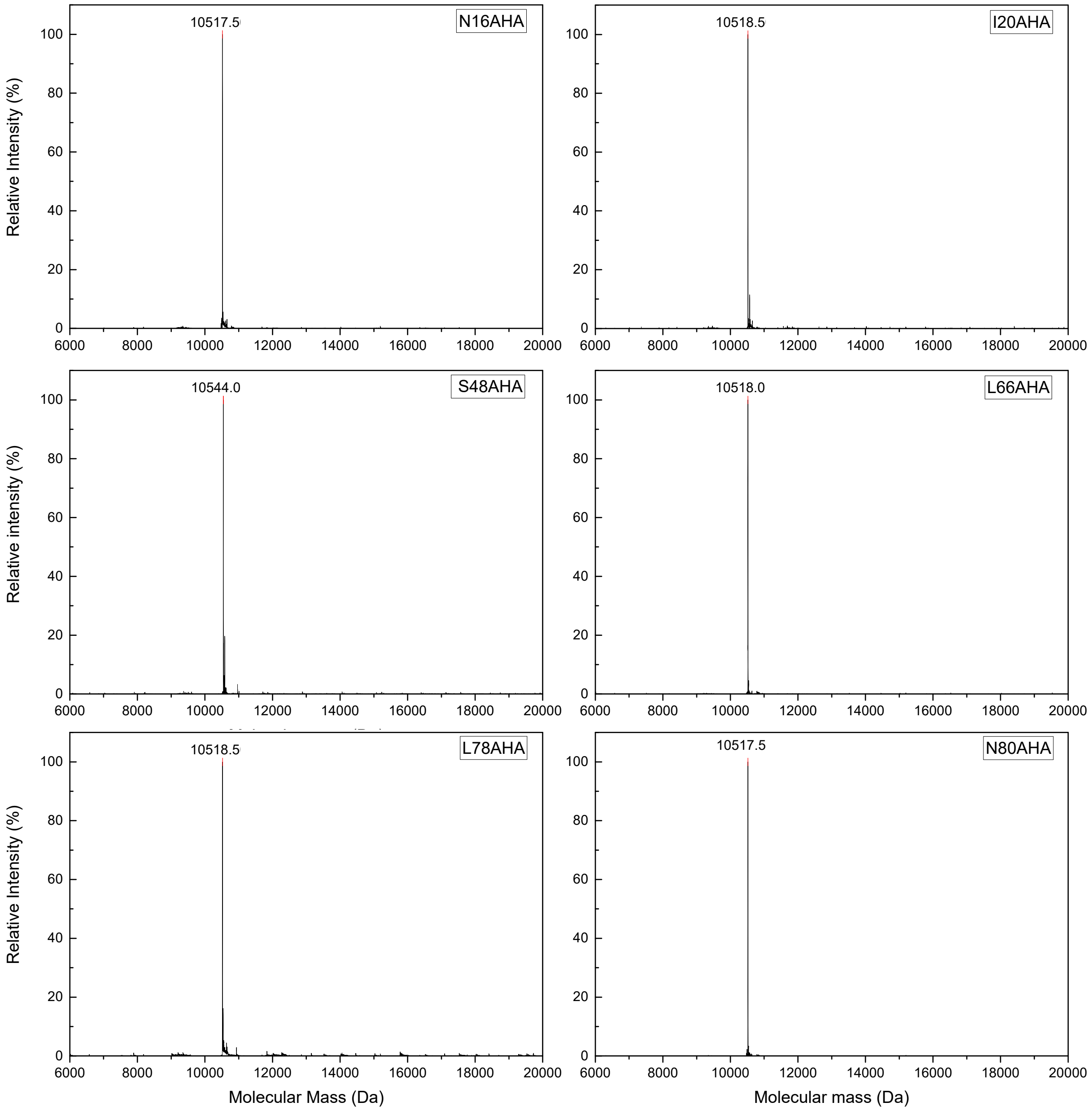


FIG. S1: ESI-mass spectra of all mutants considered in this paper, evidencing the the excellent purity of the labelled and cross-linked protein.

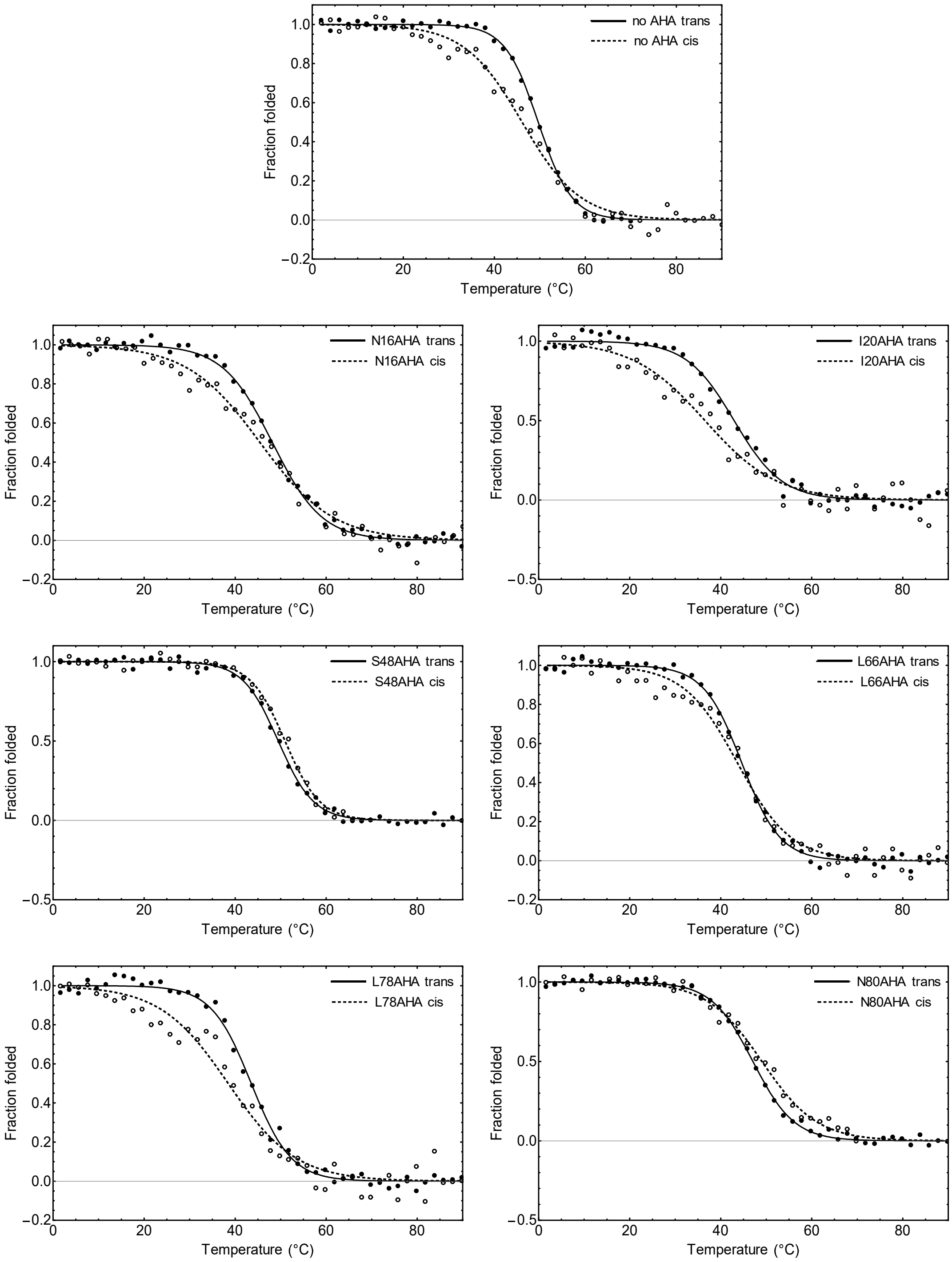


FIG. S2: Temperature-induced unfolding measured by CD in the two states of the photo-switch of the PDZ2 domain without AHA label (top), and all the mutants considered in this paper. The *trans* data have been measured for the dark-adapted protein, the *cis* data after illumination for 2 min at 370 nm from a cw diode laser (CrystaLaser CL-2000). The data have been fit to a function $1/(1+\exp((T-T_m)/\Delta T))$ after subtraction of the background and normalisation. The fit results are summarized in Table 1.

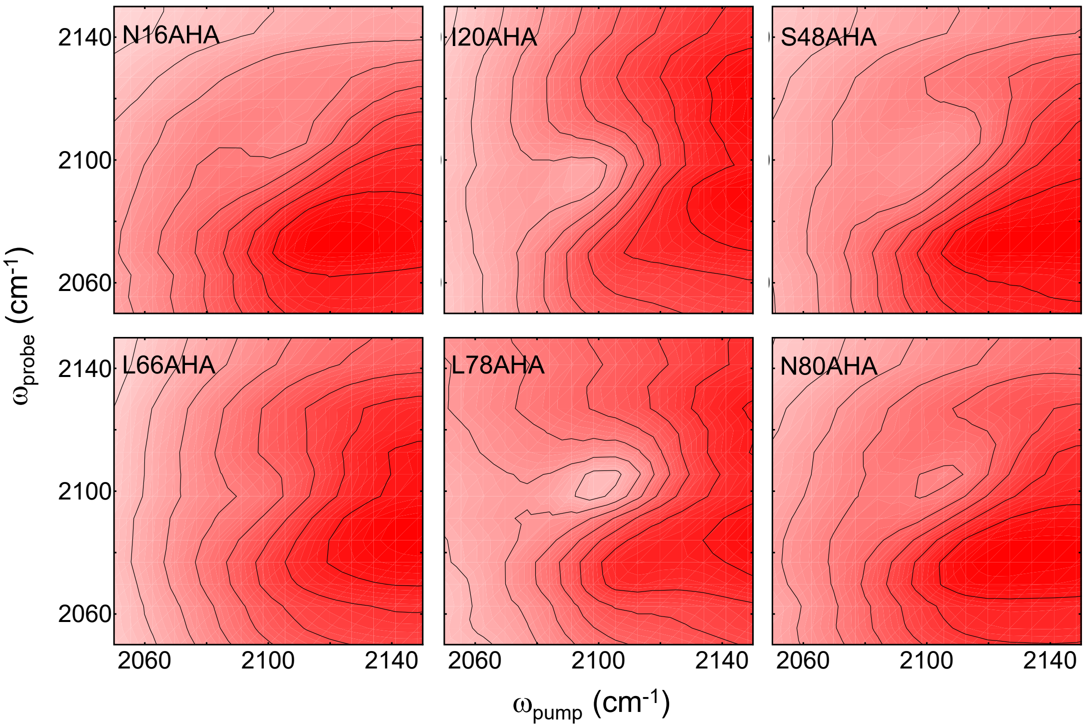


FIG. S3: Purely absorptive 2D IR spectra of the dark-adapted *trans*-state of all mutants considered in this study. The AHA signal can be recognized, but is sitting on a huge water background (which has not been measured independently and subtracted out, since that would require much larger sample amounts). The water background is positive (red, excited state absorption), while the AHA diagonal band is seen as negative dip. The relative contribution of the AHA varies, since the concentration varied between 0.6 mM and 1.1 mM, given by the solubility of the particular mutant.

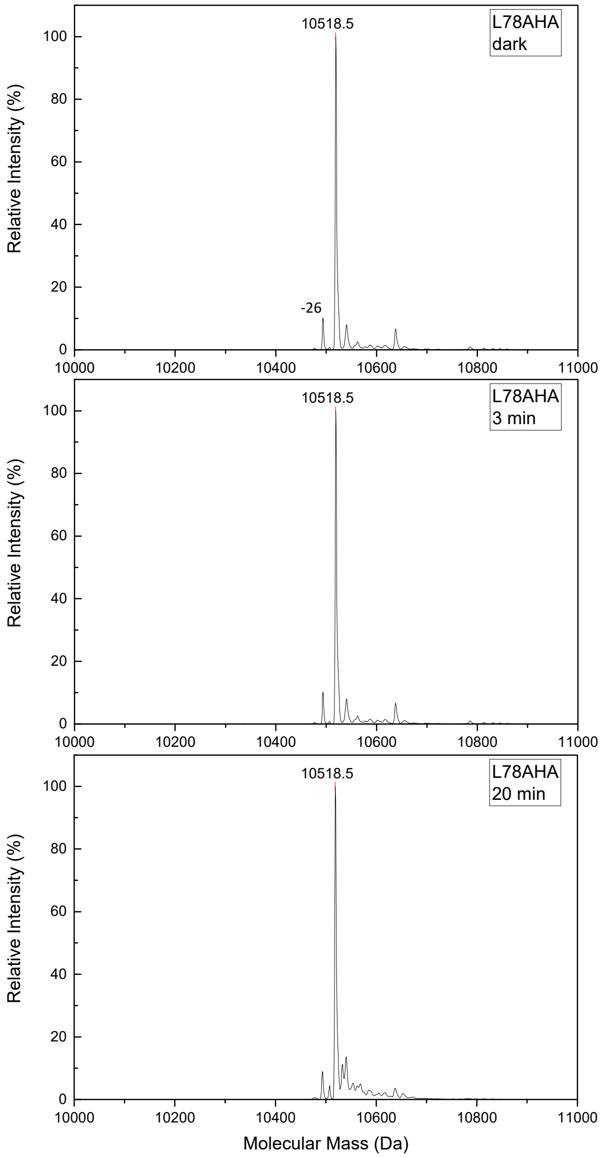


FIG. S4: ESI-mass spectra of L78AHA (from a different batch than in Fig. S1) before illumination (top), and after 3 min (middle) or 20 min (bottom) of illumination at 370 nm from a cw diode laser (CrystaLaser CL-2000). The first case coincides with the conditions used to measure the 2D IR difference spectra of Fig. 3, while the second case applied significantly more light. Even in the second case, still basically no changes in the chemical composition of the sample are detectable, in particular no reduction of the azido-group that would lower the mass by 26 Da due to the loss of a N_2 molecule. The corresponding peak is labelled in the top panel, is very small, and does not increase upon illumination.

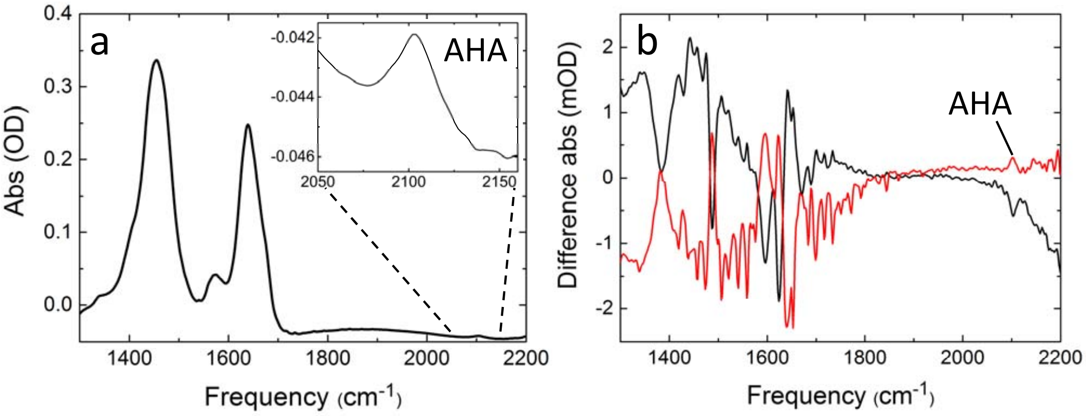


FIG. S5: (a) Absolute FTIR spectrum of L78AHA with the buffer background subtracted; the inset focusing into the small AHA band. (b) Difference FTIR spectra of L78AHA. The black line shows the result when taking a background for the dark-adapted protein in *trans*, and then switching to *cis* by illumination at 370 nm. The dip at ≈2100 cm^{-1} indicates the loss of intensity of the AHA band. The red line shows the result when taking the *cis* configuration as background and then (partially) switching back to *trans* by illumination at 420 nm. The AHA band regains its intensity. Back-switching from *cis* to *trans* is not complete, since both the *cis* and the *trans* configuration absorb at 420 nm, thus establishing a photo-equilibrium (in contrast to *trans*-*cis* switching at 370 nm that can be performed to almost 100%). The results nevertheless show that back-switching is reversible.

References:

[1] Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field, *J. Comput. Chem.* **2004**, *25*, 1157–1174.

[2] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, .; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09;* Gaussian Inc. Wallingford CT, 2009.

[3] Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model, *J. Phys. Chem.* **1993**, *97*, 10269–10280.

Chapter 3

Azidohomoalanine: a Minimally Invasive, Versatile, and Sensitive Infrared Label in Proteins to Study Ligand Binding

After mimicking the PDZ2 domain allosteric transition by photoswitching [145], observing the 2D-IR responses of six azidohomoalanine (Aha) mutants where the Aha label was placed each time in a different position within the protein structure [101], a direct observation of ligand recognition was attempted. The method was pushed to the limits; a specific Aha mutant was expressed (K38Aha), where Aha was placed at the margin of the PDZ2 binding pocket to perturb as little as possible the binding event, but in a position close enough to observe it [181] and involved in the recognition mechanism *via* salt bridge formation [182]. These peculiarities cause small but evident vibrational frequency shifts of the Aha label upon peptide binding of *circa* 1-3 cm^{-1} [35]. The photoswitch was used in this case (a water-insoluble version without the sulfonate groups ($—SO_3{}^-$) [143], see the chemical structure on the top in Fig. 1.6) to produce a peptide variant of the wild type version, which binds and affects the Aha local environment in a different manner. These small differences were detected by 2D-IR spectroscopy and molecular dynamics (MD) simulations gave a qualitative explanation for them.

- ∗ I contributed to this paper with the acquisition of the purely absorptive buffer-subtracted 2D-IR spectra in Fig. 2, with the plots in Fig. S2, and with the paper writing which has been revised by Prof. Dr. Peter Hamm.

THE JOURNAL OF PHYSICAL CHEMISTRY B
Article
Cite This: J. Phys. Chem. B 2018, 122, 10118−10125
pubs.acs.org/JPCB

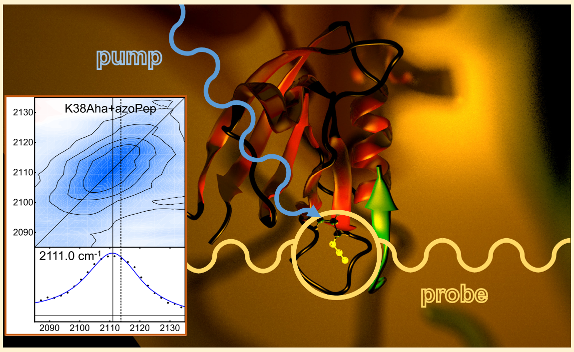
Azidohomoalanine: A Minimally Invasive, Versatile, and Sensitive Infrared Label in Proteins To Study Ligand Binding

Claudio Zanobini,[†] Olga Bozovic,[†] Brankica Jankovic,[†] Klemens L. Koziol,[†] Philip J. M. Johnson,[†] Peter Hamm,[*,†] Adnan Gulzar,[‡] Steffen Wolf,[‡] and Gerhard Stock[*,‡]

[†]Department of Chemistry, University of Zurich, Zurich 8057, Switzerland
[‡]Biomolecular Dynamics, Institute of Physics, Albert Ludwigs University, Freiburg 79104, Germany

Supporting Information

ABSTRACT: The noncanonical amino acid azidohomoalanine (Aha) is known to be an environment-sensitive infrared probe for the site-specific investigation of protein structure and dynamics. Here, the capability of that label is explored to detect protein−ligand interactions by incorporating it in the vicinity of the binding groove of a PDZ2 domain. Circular dichroism and isothermal titration calorimetry measurements reveal that the perturbation of the protein system by mutation is negligible, with minimal influence on protein stability and binding affinity. Two-dimensional infrared spectra exhibit small (1−3 cm^{-1}) but clearly measurable red shifts of the Aha vibrational frequency upon binding of two different peptide ligands, while accompanying molecular dynamics simulations suggest that these red shifts are induced by polar contacts with side chains of the peptide ligands. Hence, Aha is a versatile and minimally invasive vibrational label that is not only able to report on large structural changes during, e.g., protein folding, but also on very subtle changes of the electrostatic environment upon ligand binding.



I. INTRODUCTION

Proteins are key mediators in virtually every biological process exhibiting their functions through specific interactions with other proteins, peptides, nucleic acids, small molecules, or ions.[1−3] Vital information in the cell is transmitted through specific non-covalent interactions in e.g., signal transduction cascades, regulation of metabolic processes, activation or inhibition of enzymatic reactions, assembly of macromolecular complexes, and programmed cell death.[4,5] Disruption of specific interactions is a major cause of numerous diseases. Therefore, understanding structural details of protein−ligand interactions is a prerequisite for clarifying cellular processes at a molecular level.[6] Furthermore, knowledge of the mechanisms of protein−ligand recognition facilitates rational drug design for treatments of different diseases.[7−9] Because of its importance, a plethora of methods for protein−ligand binding analysis have been developed.[10−12] Specificity is one of the crucial features of protein−ligand interactions, and therefore, methods that can directly discriminate between specific and non-specific binding are advantageous.[13,14]

It has recently been demonstrated that the presence of an infrared label at the site of recognition between a protein and its binding partner can give site-specific insight into the underlying mechanisms of how signaling proteins function.[15−19] Such studies require a special vibrational label that absorbs in the spectral window between ≈1700 and ≈2800 cm^{-1} to discriminate it from a huge protein background.[20−25] Among the possible molecular groups that have been proposed in this regard,[16,17,19,26−44] we favor the noncanonical amino acid azidohomoalanine (Aha). Aha has a relatively high extinction coefficient (≈300−400 M^{-1} cm^{-1}),[15] and it has been shown to be an environment-sensitive infrared probe, capable of sensing the polarity of its environment via the frequency of its vibrational transition.[15,18,28,45−47] Additionally, this noncanonical amino acid can be incorporated into a protein in virtually any position via a methionine-auxotrophic expression strategy.[48−50] All these characteristics make Aha one of the most promising IR labels.

For the most part, Aha has only been used to detect quite significant structural changes in the protein structure. For example, unfolding of an Aha mutant has been investigated, where the effect has been the largest when the label is incorporated into the core of the protein, with a frequency blue shift of 19 cm^{-1} upon thermal denaturation.[45] In this case, not only the structure of the protein changes dramatically upon unfolding, but the environment of the Aha label also changes from hydrophobic in the core of a protein to fully solvent exposed. With regard to ligand binding, on the other hand, the Aha label has been incorporated into peptide fragments derived from the Ras-associated guanine nucleotide exchange

Received: August 28, 2018
Revised: October 17, 2018
Published: October 21, 2018

DOI: 10.1021/acs.jpcb.8b08368
J. Phys. Chem. B 2018, 122, 10118−10125

factor 2 (Ra-GEF2),[15,18] whose C-terminal domain interacts with the PDZ2 domain of human phosphatase 1E.[51] PDZ-domain containing proteins mediate a wide variety of signaling processes in diverse organisms, and thus represent excellent model systems for peptide binding studies.[52] Upon binding to a PDZ2 domain, the peptide ligand undergoes a structural change that is comparable to protein folding, that is, the peptide is a solvent-exposed random coil in its unbound state, but structurally well defined in its bound state. Furthermore, depending on the position of the Aha label, it points into the rather hydrophobic binding pocket, and consequently, a frequency shift of similar size (15 cm^{-1}) has been observed.[18]

With the present paper, we set out to explore how far we can push the method. That is, rather than having the Aha label in the peptide ligand,[15,18] we incorporate it in the protein sequence near the binding pocket at position K38. This site was chosen since previous molecular dynamics (MD) simulations have suggested frequent salt-bridges of the peptide ligand with K38 during the process of binding.[53] Hence, we expect at most a marginal structural change of the protein upon ligand binding,[54] but the ligand will still modulate the degree of solvation of the Aha label. We apply two-dimensional infrared (2D IR) spectroscopy[55] in a dedicated 2D IR spectrometer[23] to be sensitive enough to measure very small frequency shifts of the Aha vibration. For the purpose of this work, the advantage of 2D IR spectroscopy is its quadratic signal dependence on the absorption cross section, which to a significant extent suppresses the strong water background. We furthermore present molecular dynamics (MD) simulations aimed to provide microscopic information on the structural changes around the binding pocket, and the mechanisms of solvation of the Aha label.

We considered the three molecular systems shown in Figure 1 in this study. As a reference sample, we measured and MD simulated the K38Aha mutant of apo-PDZ2 (denoted as K38Aha throughout this paper, see Figure 1a). We then bound either the wild-type peptide (RWAKSEAKENEQVSAV) to its binding groove (denoted K38Aha + wtPep, see Figure 1b), or a mutant of that peptide (RWAKSEAKEC EQVSC V) with an apolar azobenzene moiety attached to the peptide via the two cysteines (denoted as K38Aha + azoPep, see Figure 1c). The azobenzene-variant has originally been developed to photoswitch the binding affinity of the peptide ligand, which will be the topic of another publication. Here, we make use of the hydrophobicity of the azobenzene moiety, which might affect the degree of solvation of the Aha label (note that in contrast to our previous studies,[50,56,57] the azobenzene moiety does not contain any $-SO_3^-$ groups). All experiments have been performed in the dark to ensure that the azobenzene moiety is in its trans configuration.

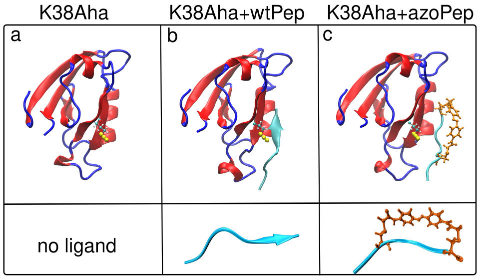


Figure 1. Molecular systems considered in this study. (a) The K38Aha mutant of apo-PDZ2 with the Aha label indicated in balls and sticks, (b) with the wild-type peptide bound to its binding groove (K38Aha + wtPep), and (c) with azobenzene-variant of the peptide bound to the binding groove (K38Aha + azoPep); azobenzene photoswitch displayed as orange sticks. The shown structures are snapshots taken from the MD simulation.

II. METHODS

II.I. Protein and Peptide Preparation. Expression of the K38Aha mutant of PDZ2 domain was performed as described earlier.[49,50] The two peptide variants, wild-type peptide (RWAKSEAKENEQVSAV) and its cysteine mutant (RWAKSEAKECEQVSCV), were synthesized using solid phase peptide synthesis with standard 9-fluorenylmethoxycarbonyl (Fmoc) chemistry on a Liberty 1 peptide synthesizer (CEM Corporation, Matthews, NC). The crude products were purified using reverse phase high-performance liquid chromatography (HPLC) with a C18 column (Macherey Nagel, Duren, Germany), 0.1% trifluoroacetic acid buffered acetonitrile gradient 0−50% in 10 column volumes.

The second peptide variant was linked to the azobenzene moiety (di-iodoacetamide azobenzene) as reported previously[58] with the following modifications: 50 μmol of peptide was dissolved in 75 mL of 20 mM Tris buffer, pH 8.5, which was previously extensively degassed. To promote the linking reaction by reducing disulfide bridges between cysteines, 1 equiv of tris(2-carboxyethyl)phosphine was added to peptide solution and allowed to incubate for 1 h after which 100 μmol of azobenzene linker dissolved in 250 mL of tetrahydrofuran was added. The solution was stirred and left in the dark at room temperature overnight. The linked peptide was concentrated at 30 °C to precipitate the excess of apolar azobenzene linker. The suspension was filtered and the linked azo-peptide was repurified using reverse phase HPLC with a C18 column.

All products, proteins and both peptides, were dialyzed against 50 mM borate, 150 mM NaCl buffer, pH 8.5, lyophilized, and finally resuspended in H_2O for the circular dichroism (CD) and isothermal titration calorimetry (ITC) measurements and D_2O for the 2D IR measurements. Purity of all samples was confirmed by mass spectrometry analysis.

II.II. Two-Dimensional IR Spectroscopy. Two-dimensional IR spectra in the boxcar geometry were measured using mid-IR pulses generated by a home-built optical parametric amplifier[59] providing ≈150 fs full width at half maximum (FWHM), ≈2 μJ pulses at ≈5 μm with a repetition rate of 5 kHz, as previously described.[18] A syringe pump sample delivery system was used to flow either the sample or buffer through a cell comprised of two CaF_2 windows (2 mm thickness, 25 mm diameter) with a 25 μm Teflon spacer. Two-dimensional IR spectra of the buffer and of the sample solutions were measured independently, exchanging approximately 250−300 μL of sample per measurement with the help of the syringe pumps. Buffer and sample spectra were phased individually, using the water background for a reference phase as described previously,[60] and then subtracted to reveal the isolated Aha signal.

II.III. Computational Methods. The starting structure of PDZ2 was taken from PDB entry 3LNX.[61] To generate initial structures for the MD simulations of K38Aha + wtPep, we extracted the peptide structure and position from the X-ray structure 3LNY,[61] which contains only a six residue long peptide ligand, and attached three additional residues from the

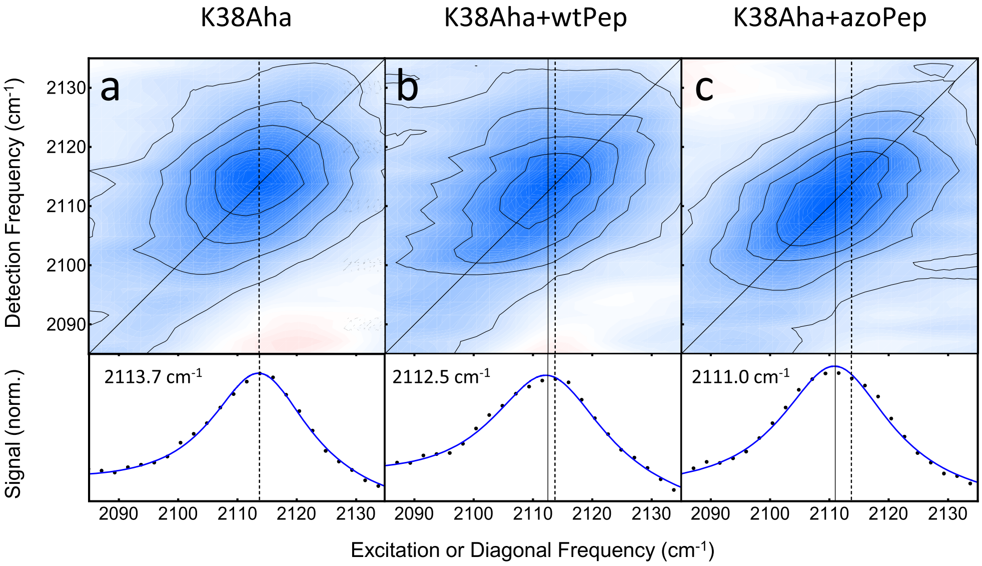


Figure 2. Ligand binding observed with 2D IR spectroscopy. (a) Two-dimensional IR response of K38Aha, (b) K38Aha + wtPep, and (c) K38Aha + azoPep. The bottom panels plot the diagonal signal together with fits used to deduce the peak position (see text for details). In panel (a) the dotted line marks the reference peak position of the Aha label without any ligand. In panels (b) + (c), the peak positions of the Aha label with the two different ligands are marked in addition to it as solid lines.

NMR structure of 1D5G[54] at the N-terminus. The peptide considered in the MD simulation (KENEQVSAV), thus, is shorter than in experiment. K38 was mutated to Aha in all systems.

MD equilibrium simulations with a production run length of 100 ns were performed using the GROMACS software package v2016.3[62] with a hybrid GPU-CPU acceleration scheme and the AMBER99SB-*ILDN forcefields as described in ref 50. Azo-switch parameters were obtained using Antechamber[63] with BCC charges[64,65] and are described in the Supporting Information. We used a dodecahedral box with an image (i.e., face-to-face) distance of 7 nm. The box contained 7495 (K38Aha), 7997 (K38Aha + wtPep), and 10 168 (K38Aha + azoPep) water molecules, respectively. Further simulation details can be found in ref 50.

To determine $-N_3$ contacts with the protein and the ligand, g_mindist from the Gromacs tools was employed. Contact distributions were then obtained by histogramming the MD data with 0.01 nm binning width. We define a contact to be formed if the minimal distance between azido group nitrogen atoms and protein/ligand residues is shorter than 0.45 nm.[66,67] In a similar way, we analyzed contacts between Aha and water as azido group/water oxygen atom distances with a cutoff of 0.45 nm.

We used the empirical model of Cho and co-workers[47] to estimate vibrational frequency shifts $\delta\omega$ caused by changes in the electrostatic environment of the Aha labels. By calculating the electric field $E_j(t)$ at the nitrogen atoms (j = 1, 2, 3) of the azido group for each MD snapshot at time t, we obtain the spectral shift (relative to the vacuum value)

$$\delta\omega(t) = \sum_j a_j E_j(t) \tag{1}$$

with coefficients a_j given in ref 47. Electric fields were computed via a reaction field approach using a cutoff radius r_c = 2.3 nm as described in ref 50. From the frequency trajectory $\delta\omega(t)$ with a time step of 15 ps, the distribution of the vibrational shifts was obtained via a histogram using 50 bins between −25 and 25 cm^{-1}.

III. RESULTS

III.I. Experimental Results. To set the stage, we start with verifying that the Aha label only minimally disturbs the protein system. To that end, temperature dependent CD signals were recorded at 205 nm, revealing melting temperatures of 44 ± 1 and 46 ± 1 °C for wild-type PDZ2 and K38Aha (Figure S1, Supporting Information), respectively. Hence, the mutation to Aha, in fact slightly stabilizes the protein. Moreover, the dissociation constants of the wild-type peptide to wild-type protein and K38Aha are virtually the same (K_D = 11 μM, see Figure S2a,b, Supporting Information). This is the evidence of the fact that the energy of binding has not been affected by substituting Aha in the vicinity of the binding groove of the protein. We also measured the binding affinity of the azo-peptide to K38Aha, revealing a somewhat smaller value (K_D = 3 μM, see Figure S2c, Supporting Information) compared to the wild-type peptide, indicating that the azo-moiety points away from the binding groove when bound to the protein, but capable of interacting with the protein in a stabilizing manner.

Figure 2, top panels, show 2D IR spectra of the Aha vibrational label of all three sample systems (we concentrate here on the ground-state (0−1) peak of the 2D IR response on the diagonal, while the excited state (1−2) peak is outside the frequency window shown in Figure 2). The concentrations used in these experiments were 0.9 mM for the protein, and 2.5 and 1.5 mM for wild-type and the azo-peptide, respectively. Given the binding affinities of both peptides in the low micromolar range (see the Supporting Information, Figure S2), practically 100% of the protein, which carries the Aha label, has a ligand bound at these concentrations. Figure 2, bottom panels show diagonal cuts through the 2D IR spectra (averaging over the diagonal and the two first off-diagonals). To determine the peak frequency of the label, we fit these diagonal cuts to a lineshape function. We found that the

DOI: 10.1021/acs.jpcb.8b08368
J. Phys. Chem. B 2018, 122, 10118−10125

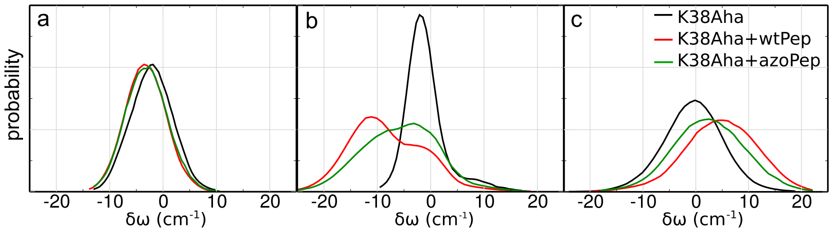


Figure 3. Vibrational frequency shifts (relative to the vacuum value) of all simulated systems by considering (a) all atoms, (b) only protein/ligand atoms and (c) only solvent atoms. The black curve shows the result for K38Aha, and the red and green curves that of K38Aha + wtPep and K38Aha + azoPep, respectively.

quality of the fit is better when using a Lorentzian rather than a Gaussian lineshape, and we added a tilted linear baseline to the fit, which accounts for a residual water background. We considered the center of the fitted Lorentzian to be the peak position of the corresponding peak band. This procedure revealed a peak position of 2113.7 cm^{-1} for K38Aha (marked as dotted vertical lines in Figure 2a−c), and red shifted peaks at 2112.5 cm^{-1} for K38Aha + wtPep and 2111.0 cm^{-1} for K38Aha + azoPep (solid vertical lines in Figure 2b,c). In K38Aha, the linewidth is 20.5 cm^{-1} FWHM together with a rather round 2D IR lineshape (justifying the Lorentzian fit). Upon binding of a ligand, the linewidth increases a little bit (23.6 cm^{-1} for K38Aha + wtPep and 22.6 cm^{-1} for K38Aha + azoPep) and the 2D IR lineshape becomes more elongated along the diagonal, indicating a more inhomogeneously broadened absorption band.

By repeating these experiments various times (i.e., between 2 and 4 times for different samples), taking into account the uncertainty in the phasing,[60] and by fitting the results to different models (e.g., fitting only one-dimensional diagonal cuts as in Figure 2 vs fitting the complete 2D lineshape or using different lineshape functions), we estimated that the statistical and systematic error in the determination of the peak position of the Aha band is in the order of 0.5 cm^{-1}. One example of a completely different data set of K38Aha and K38Aha + wtPep is shown in Figure S3, which has been measured with different concentrations on different days by different people. Applying the same fit function, the deduced peak frequencies are the same within 0.1 cm^{-1} as shown in Figure 2, evidencing the reproducibility of the measurement. To conclude this part, for both K38Aha + wtPep and K38Aha + azoPep, we are confident that the effect is real with frequency red shifts of 1.2 ± 0.5 and 2.7 ± 0.5 cm^{-1}, respectively.

III.II. Computational Results. Recent quantum-chemical calculations of Cho and co-workers[46,47] have shown that the spectroscopic signatures of the azido stretch mode of Aha mainly reflect the local electrostatic environment of the $-N_3$ reporter group. The structures shown in Figure 1 suggest that the azido group is indeed found in different environments in each investigated case. That is, Aha is fully exposed to the solvent in K38Aha, and rarely interacts with protein atoms, while in the cases of K38Aha + wtPep and K38Aha + azoPep it is less exposed to the solvent, but interacts mostly with ligand residues. These differences are expected to be detectable in the vibrational spectra.

Indeed, using eq 1 for calculating spectral shifts displayed in Figure 3a, we found that the reporter group in K38Aha + wtPep and K38Aha + azoPep induces a red shift of 1−2 cm^{-1} with respect to K38Aha. To learn about the contribution of protein and solvent to the total vibrational frequency, we calculated the shifts due to protein/ligand atoms and solvent atoms independently. Considering only protein/ligand atoms, K38Aha + wtPep and K38Aha + azoPep show red shifts of 8−10 and 1−3 cm^{-1} compared to K38Aha (Figure 3b). Regarding solvent molecules only, K38Aha + wtPep and K38Aha + azoPep are blue shifted by 2−5 cm^{-1} (Figure 3c).

To relate the vibrational shifts to changes in contact patterns, Figure 4 compares the probability distribution of the number of azido contacts with protein/ligand and water molecules. Concerning protein/ligand contacts, K38Aha + wtPep and K38Aha + azoPep simulations exhibit 12−18 protein/ligand contacts compared to K38Aha with nine contacts. Regarding water contacts, the azido group of K38Aha + wtPep and K38Aha + azoPep exhibits 7−8 water contacts, while K38Aha on average forms 10−12 contacts with water molecules. Hence, water contacts are replaced by ligand contacts upon binding of the peptide ligand.

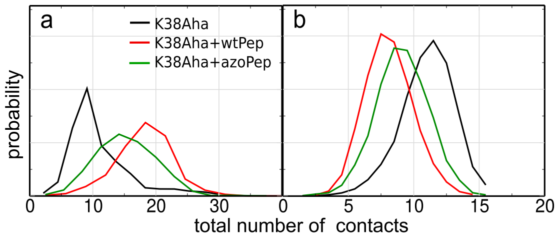


Figure 4. Probability distribution of average number of contacts of the azido group with (a) protein/ligand and (b) solvent molecules. The black curve shows the result for K38Aha, and the red and green curves that of K38Aha + wtPep and K38Aha + azoPep, respectively.

To further explore the cause of red shifts of K38Aha + wtPep and K38Aha + azoPep, we calculated minimum distances between the azido group and all residues of the respective systems. We found that significant contacts only appear with ligand residues (Figure S4), as suggested in ref 53. From the minimum distance analysis, three major contacts of Aha38 with ligand residues Glu(−5), Gln(−4), and Val(−3) were identified as potential candidates to induce red shifts. However, the distance patterns of Gln(−4) and Val(−3) do not significantly change between both protein/ligand systems, while Glu(−5) is special, as it is the only one of the three residues with significant changes in distances. That is, the middle nitrogen atom of the azido group of K38Aha + wtPep forms a consistent polar contact with the negatively charged side chain carboxyl oxygen of Glu(−5) (Figure 5a), while this contact is less stable in K38Aha + azoPep. To test if the contact with Glu(−5) is the main origin of red shifts observed in experiments and simulations, we excluded this specific

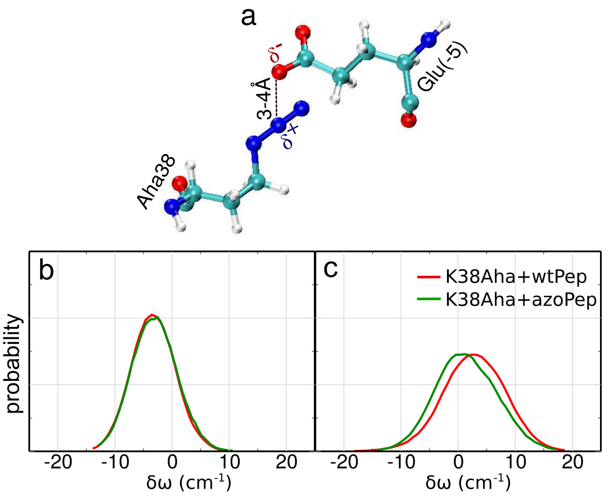


Figure 5. (a) Polar contact between the positively charged $N^{(2)}$ of the azido group and the negatively charged side chain carboxyl oxygen of Glu(−5). (b, c) Vibrational frequency shifts of K38Aha + wtPep and K38Aha + azoPep by including (b) and excluding (c) Glu(−5). The red curve shows the result for K38Aha + wtPep and the green that for K38Aha + azoPep.

residue and compared the resulting spectrum with spectra including Glu(−5) (Figure 5b,c). Exclusion of Glu(−5) induces a blue shift of ~4−6 cm^{-1}, with a difference of ~2−3 cm^{-1} between K38Aha + wtPep and K38Aha + azoPep. Very likely, it is these very specific, but fluctuating contacts that give rise to the larger inhomogeneity observed experimentally upon ligand binding in the form of 2D IR lineshapes that are more elongated along the diagonal (Figure 2b,c).

We also looked into the possible contribution of the ions in the solution to the frequency shift. Figure S5 shows that the average distance between the azido group to the nearest ion is between 1.0 and 1.5 nm, and contacts (i.e., distances of less than 4.5 Å) almost never occur. We therefore conclude that the effect of the ions on vibrational frequency of the azido group is negligible due to the shielding by water.

IV. DISCUSSION AND CONCLUSIONS

We have demonstrated that it is possible to reliably measure frequency shifts of the Aha label as small as ≈1 cm^{-1} with the help of 2D IR spectroscopy, sufficient to detect rather subtle structural changes of a protein system. Furthermore, MD simulation together with pre-parameterized spectroscopic maps[47] can qualitatively reproduce the size of the spectroscopic effect. However, while the experiments reveal a bigger effect for K38Aha + azoPep than for K38Aha + wtPep, consistent with the notion that the hydrophobic azobenzene moiety might shield the Aha label more efficiently from the solvent, that effect is not reproduced by the MD simulations. On the contrary, the MD simulations suggest that interactions with very specific side chains (in this case Glu(−5)) might have a dominating effect, the exact calculation of which would require more sophisticated quantum mechanics/molecular mechanics simulations.[47]

It is the common notion that the frequency shift of the azido group is a measure of the amount of solvation; the stronger an azido group is solvated, the more blue is its absorption band.[15,18,28,38,45−47] Our experimental observation supports this view, with the ligands shielding the Aha label to a certain extent from water (Figure 2). The MD results of Figure 4b indeed show that the number of solvent contacts is reduced once a ligand is bound to the protein. Yet, Figures 3, 4a, and 5 suggest that this common notion needs to be re-interpreted somewhat. That is, a smaller number of water contacts is inevitably compensated by a larger number of protein contacts (Figure 4a vs 4b), which can be polar as well; in this case, in particular, due to Glu(−5) from the ligand (Figure 5). These protein contacts also reveal frequency shifts due to electrostatic interactions according to eq 1. In fact, since the protein is more structured than the solvent, these protein contacts are more stable on average and may overcompensate the contribution of water in terms of the frequency shift (Figure 3). In other words, Figure 3 reveals that all polar interactions tend to produce a frequency red shift, in contrast to the common notion, and that reduced solvation by water is overcompensated by even stronger "solvation" from the protein. It should, however, be noted that this interpretation is at odds with the conclusions of ref 38, which measured the frequency of isolated azido-groups also in very apolar solvents such as hexane, which might be assumed to be close to the situation in vacuum, but where the frequency is red shifted relative to that in water. However, in the present case, assuming the protein to be a hydrophobic bulk turns out not to be a viable approximation. Instead, especially due to the presence of Glu(−5), the protein resembles best a highly anisotropic polar "solvent" for the Aha azido group with stable charge distribution, resulting in the observed additional red shifts.

It is interesting to compare the present results to those of previous work, where we used Aha labels at various positions to detect structural changes in a PDZ2 domain that has been made photo-switchable by covalently linking a similar azo-moiety across its binding groove.[50] Only L78Aha, which was present very close to the photoswitch in the binding groove, showed an effect; in this case, a change in the absorption intensity rather than frequency. Also, for this example we concluded that the effect is related to very specific interactions of the Aha label with the azo-photoswitch.

Previous work[15,18] has shown that Aha can be used as a reporter to determine biomolecular recognition when incorporated into the peptide ligand directly. In this study, Aha has instead been incorporated into the protein in the vicinity of the binding groove. In contrast to the ligand-incorporated reporter, we have shown by CD and ITC experiments that the presence of the Aha label neither perturbs the stability of the PDZ2 domain, nor its binding affinity for the peptide ligand, in agreement with recent NMR and X-ray experiments that came to the same conclusion for a similar molecular system.[68] The 2D IR spectra of Figure 2 showed that Aha is indeed a sensitive probe of ligand binding, exhibiting a few wavenumbers red shift upon binding of two different peptides. MD simulations can reproduce the effect essentially quantitatively, providing an atomistic picture of the interactions of the Aha label at the protein surface. It is important to stress that the center frequency of an isolated vibrational transitions, such as that of the Aha label, can be measured with an accuracy much better than its linewidth or than the spectral resolution of the 2D IR instrument. This is the equivalent of the very idea of sub-diffraction microscopy, where the position of individual switchable fluorescing chromophores can be measured with a precision much better than the resolution of the microscope, provided that their point spread functions do not spatially overlap.[69] In conclusion, the present study demonstrates that

Aha can be employed as a specific IR reporter not only for big changes of its chemical environment (e.g., protein folding/unfolding events), but also for very subtle changes of the electrostatic environment at the protein surface.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.8b08368.

> CD data (Figure S1), ITC data (Figure S2), completely independently measured set of 2D IR (Figure S3), contact analysis of the azido group with the ligand residue (Figure S4) and the ions in solution (Figure S5), Gromacs parameters for the azo-moiety (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: peter.hamm@chem.uzh.ch (P.H.).
*E-mail: stock@physik.uni-freiburg.de (G.S.).

ORCID

Philip J. M. Johnson: 0000-0002-7251-4815
Peter Hamm: 0000-0003-1106-6032
Steffen Wolf: 0000-0003-1752-6175
Gerhard Stock: 0000-0002-3302-3044

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Ben Schuler and his group for their continuous and tremendous help with the protein chemistry, Amedeo Caflisch for many insightful discussions regarding ligand binding to the PDZ2 domain, Rolf Pfister for the synthesis of the peptides, and the Functional Genomics Center Zurich, especially Serge Chesnov, for help with the mass spectrometry. The work has been supported in part by the Swiss National Science Foundation (SNF) through the NCCR MUST and Grant 200021_165789/1, as well as by the Deutsche Forschungsgemeinschaft through Grant STO 247/10-1. Computational resources were provided by the bwForCluster BinAC (RV bw16I016) and the Black Forest Grid Initiative.

■ REFERENCES

(1) Pang, X.; Zhou, H.-X. Rate constants and mechanisms of protein-ligand binding. *Annu. Rev. Biophys.* **2017**, *46*, 105−130.

(2) Stelzl, U.; Worm, U.; Lalowski, M.; Haenig, C.; Brembeck, F. H.; Goehler, H.; Stroedicke, M.; Zenkner, M.; Schoenherr, A.; Koeppen, S.; et al. A human protein-protein interaction network: A resource for annotating the proteome. *Cell* **2005**, *122*, 957−968.

(3) Stanfield, R. L.; Wilson, I. A. Protein-peptide interactions. *Curr. Opin. Struct. Biol.* **1995**, *5*, 103−113.

(4) McCammon, J. A. Theory of biomolecular recognition. *Curr. Opin. Struct. Biol.* **1998**, *8*, 245−249.

(5) Huang, Z. The chemical biology of apoptosis: Exploring protein-protein interactions and the life and death of cells with small molecules. *Chem. Biol.* **2002**, *9*, 1059−1072.

(6) Arkin, M. R.; Tang, Y.; Wells, J. A. Small-molecule inhibitors of protein-protein interactions: profressing towards the reality. *Chem. Biol.* **2014**, *21*, 1102−1114.

(7) Babine, R. E.; Bender, S. L. Molecular recognition of protein-ligand complexes: Applications to drug design. *Chem. Rev.* **1997**, *97*, 1359−1472.

(8) Kastritis, P. L.; Bonvin, A. M. J. J. On the binding affinity of macromolecular interactions: daring to ask why proteins interact. *J. R. Soc., Interface* **2013**, *10*, No. 20120835.

(9) Zhao, L.; Chmielewski, J. Inhibiting protein-protein interactions using designed molecules. *Curr. Opin. Struct. Biol.* **2005**, *15*, 31−34.

(10) Piehler, J. New methodologies for measuring protein interactions in vivo and in vitro. *Curr. Opin. Struct. Biol.* **2005**, *15*, 4−14.

(11) *Protein-Ligand Interactions Methods and Applications*; Williams, M. A.; Daviter, T., Eds.; Springer: New York, 2013.

(12) Greenfield, N. J. Applications of circular dichroism in protein and peptide analysis. *TrAC, Trends Anal. Chem.* **1999**, *18*, 236−244.

(13) Berggård, T.; Linse, S.; James, P. Methods for the detection and analysis of protein-protein interactions. *Proteomics* **2007**, *7*, 2833−2842.

(14) Yan, Z.; Wang, J. Specificity quantification of biomolecular recognition and its implication for drug discovery. *Sci. Rep.* **2012**, *2*, No. 309.

(15) Bloem, R.; Koziol, K.; Waldauer, S.; Buchli, B.; Walser, R.; Samatanga, B.; Jelesarov, I.; Hamm, P. Ligand binding studied by 2D IR spectroscopy using the azidohomoalanine label. *J. Phys. Chem. B* **2012**, *116*, 13705−13712.

(16) Horness, R. E.; Basom, E. J.; Thielges, M. C. Site-selective characterization of Src homology 3 domain molecular recognition with cyanophenylalanine infrared probes. *Anal. Methods* **2015**, *7*, 7234−7241.

(17) Horness, R. E.; Basom, E. J.; Mayer, J. P.; Thielges, M. C. Resolution of site-specific conformational heterogeneity in proline-rich molecular recognition by Src homology 3 domains. *J. Am. Chem. Soc.* **2016**, *138*, 1130−1133.

(18) Johnson, P. J. M.; Koziol, K. L.; Hamm, P. Quantifying biomolecular recognition with site-specific 2D infrared probes. *J. Phys. Chem. Lett.* **2017**, *8*, 2280−2284.

(19) Basom, E. J.; Manifold, B. A.; Thielges, M. C. Conformational heterogeneity and the affinity of substrate molecular recognition by cytochrome P450cam. *Biochemistry* **2017**, *56*, 3248−3256.

(20) Waegele, M. M.; Culik, R. M.; Gai, F. Site-specific spectroscopic reporters of the local electric field, hydration, structure, and dynamics of biomolecules. *J. Phys. Chem. Lett.* **2011**, *2*, 2598−2609.

(21) Adhikary, R.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. IR probes of protein microenvironments: utility and potential for perturbation. *ChemPhysChem* **2014**, *15*, 849−853.

(22) Kim, H.; Cho, M. Infrared probes for studying the structure and dynamics of biomolecules. *Chem. Rev.* **2013**, *113*, 5817−5847.

(23) Koziol, K. L.; Johnson, P. J. M.; Stucki-Buchli, B.; Waldauer, S. A.; Hamm, P. Fast infrared spectroscopy of protein dynamics: Advancing sensitivity and selectivity. *Curr. Opin. Struct. Biol.* **2015**, *34*, 1−6.

(24) Ma, J.; Pazos, I. M.; Zhang, W.; Culik, R. M.; Gai, F. Site-specific infrared probes of proteins. *Annu. Rev. Phys. Chem.* **2015**, *66*, 357−377.

(25) Błasiak, B.; Londergan, C. H.; Webb, L. J.; Cho, M. Vibrational probes: From small molecule solvatochromism theory and experiments to applications in complex systems. *Acc. Chem. Res.* **2017**, *50*, 968−976.

(26) Getahun, Z.; Huang, C. Y.; Wang, T.; DeLeon, B.; DeGrado, W. F.; Gai, F. Using nitrile-derivatized amino acids as infrared probes of local environment. *J. Am. Chem. Soc.* **2003**, *125*, 405−411.

(27) Koziński, M.; Garrett-Roe, S.; Hamm, P. 2D-IR Spectroscopy of the sulfhydryl band of cysteines in the hydrophobic core of proteins. *J. Phys. Chem. B* **2008**, *112*, 7645−7650.

(28) Oh, K.-I.; Lee, J.-H.; Joo, C.; Han, H.; Cho, M. Beta-azidoalanine as an IR probe: application fo amyloid Abeta(16-22) aggregation. *J. Phys. Chem. B* **2008**, *112*, 10352−10357.

(29) Naraharisetty, S. R. G.; Kasyanenko, V. M.; Zimmermann, J.; Thielges, M.; Romesberg, F. E.; Rubtsov, I. V. C-D Modes of deuterated side chain of leucine as structural reporters via dual-

frequency two-dimensional infrared spectroscopy. *J. Phys. Chem. B* **2009**, *113*, 4940−4946.

(30) Zimmermann, J.; Thielges, M. C.; Yu, W.; Dawson, P. E.; Romesberg, F. E. Carbon-deuterium bonds as site-specific and nonperturbative probes for time-resolved studies of protein dynamics and folding. *J. Phys. Chem. Lett.* **2011**, *2*, 412−416.

(31) Thielges, M. C.; Axup, J. Y.; Wong, D.; Lee, H. S.; Chung, J. K.; Schultz, P. G.; Fayer, M. D. Two-dimensional IR spectroscopy of protein dynamics using two vibrational labels: A site-specific genetically encoded unnatural amino acid and an active site ligand. *J. Phys. Chem. B* **2011**, *115*, 11294−11304.

(32) Thielges, M. C.; Fayer, M. D. Protein dynamics studied with ultrafast two-dimensional infrared vibrational echo spectroscopy. *Acc. Chem. Res.* **2012**, *45*, 1866−1874.

(33) Woys, A. M.; Mukherjee, S. S.; Skoff, D. R.; Moran, S. D.; Zanni, M. T. A strongly absorbing class of non-natural labels for probing protein electrostatics and solvation with FTIR and 2D IR spectroscopies. *J. Phys. Chem. B* **2013**, *117*, 5009−5018.

(34) Peran, I.; Oudenhoven, T.; Woys, A. M.; Watson, M. D.; Zhang, T. O.; Carrico, I.; Zanni, M. T.; Raleigh, D. P. General strategy for the bioorthogonal incorporation of strongly absorbing, solvation-sensitive infrared probes into proteins. *J. Phys. Chem. B* **2014**, *118*, 7946−7953.

(35) King, J. T.; Arthur, E. J.; Brooks, C. L.; Kubarych, K. J. Crowding induced collective hydration of biological macromolecules over extended distances. *J. Am. Chem. Soc.* **2014**, *136*, 188−194.

(36) Bagchi, S.; Boxer, S. G.; Fayer, M. D. Ribonuclease S dynamics measured using a nitrile label with 2D IR vibrational echo spectroscopy. *J. Phys. Chem. B* **2012**, *116*, 4034−4042.

(37) Zimmermann, J.; Thielges, M. C.; Seo, Y. J.; Dawson, P. E.; Romesberg, F. E. Cyano groups as probes of protein microenvironments and dynamics. *Angew. Chem., Int. Ed.* **2011**, *50*, 8333−8337.

(38) Wolfshorndl, M. P.; Baskin, R.; Dhawan, I.; Londergan, C. H. Covalently bound azido groups are very specific water sensors, even in hydrogen-bonding environments. *J. Phys. Chem. B* **2012**, *116*, 1172−1179.

(39) Pazos, I. M.; Ghosh, A.; Tucker, M. J.; Gai, F. Ester carbonyl vibration as a sensitive probe of protein local electric field. *Angew. Chem., Int. Ed.* **2014**, *53*, 6080−6084.

(40) van Wilderen, L. J. G. W.; Kern-Michler, D.; Müller-Werkmeister, H. M.; Bredenbeck, J. Vibrational dynamics and solvatochromism of the label SCN in various solvents and hemoglobin by time dependent IR and 2D-IR spectroscopy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 19643−19653.

(41) Le Sueur, A. L.; Schaugaard, R. N.; Baik, M. H.; Thielges, M. C. Methionine ligand interaction in a blue copper protein characterized by site-selective infrared spectroscopy. *J. Am. Chem. Soc.* **2016**, *138*, 7187−7193.

(42) Basom, E. J.; Maj, M.; Cho, M.; Thielges, M. C. Site-specific characterization of cytochrome P450cam conformations by infrared spectroscopy. *Anal. Chem.* **2016**, *88*, 6598−6606.

(43) Xu, R. J.; Blasiak, B.; Cho, M.; Layfield, J. P.; Londergan, C. H. A direct, quantitative connection between molecular dynamics simulations and vibrational probe line shapes. *J. Phys. Chem. Lett.* **2018**, *9*, 2560−2567.

(44) Lee, G.; Kossowska, D.; Lim, J.; Kim, S.; Han, H.; Kwak, K.; Cho, M. Cyanamide as an infrared reporter: comparison of vibrational properties between nitriles bonded to N and C atoms. *J. Phys. Chem. B* **2018**, *122*, 4035−4044.

(45) Taskent-Sezgin, H.; Chung, J.; Banerjee, P. S.; Nagarajan, S.; Dyer, R. B.; Carrico, I.; Raleigh, D. P. Azidohomoalanine: A Conformationally Sensitive IR Probe of Protein Folding Protein Structure and Electrostatics. *Angew. Chem., Int. Ed.* **2010**, *49*, 7473−7475.

(46) Choi, J. H.; Raleigh, D.; Cho, M. Azido homoalanine is a useful infrared Probe for monitoring local electrostatistics and side-chain solvation in proteins. *J. Phys. Chem. Lett.* **2011**, *2*, 2158−2162.

(47) Choi, J.-H.; Oh, K.-I.; Cho, M. Azido-derivatized compounds as IR probes of local electrostatic environment: Theoretical studies. *J. Chem. Phys.* **2008**, *129*, No. 174512.

(48) Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 19.

(49) Simon, M.; Zangemeister-Wittke, U.; Plückthun, A. Facile double-functionalization of designed ankyrin repeat proteins using click and thiol chemistries. *Bioconjugate Chem.* **2012**, *23*, 279−286.

(50) Stucki-Buchli, B.; Johnson, P. J. M.; Bozovic, O.; Zanobini, C.; Koziol, K. L.; Hamm, P.; Gulzar, A.; Wolf, S.; Buchenberg, S.; Stock, G. 2D-IR spectroscopy of an AHA labelled photoswitchable PDZ2 domain. *J. Phys. Chem. A* **2017**, *121*, 9435−9445.

(51) Fuentes, E. J.; Der, C. J.; Lee, A. L. Ligand-dependent dynamics and intramolecular signaling in a PDZ Domain. *J. Mol. Biol.* **2004**, *335*, 1105−1115.

(52) Lee, H.-J.; Zheng, J. J. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun. Signaling* **2010**, *8*, No. 8.

(53) Blöchliger, N.; Xu, M.; Caflisch, A. Peptide binding to a PDZ domain by electrostatic steering via nonnative salt bridges. *Biophys. J.* **2015**, *108*, 2362−2370.

(54) Kozlov, G.; Banville, D.; Gehring, K.; Ekiel, I. Solution Structure of the PDZ2 Domain from Cytosolic Human Phosphatase hPTP1E Complexed with a Peptide Reveals Contribution of the Beta 2-Beta 3 Loop to PDZ Domain-Ligand Interactions. *J. Mol. Biol.* **2002**, *320*, 813−820.

(55) Hamm, P.; Zanni, M. T. *Concepts and Methods of 2D Infrared Spectroscopy*; Cambridge University Press: Cambridge, 2011.

(56) Buchli, B.; Waldauer, S. A.; Walser, R.; Donten, M. L.; Pfister, R.; Blöchliger, N.; Steiner, S.; Caflisch, A.; Zerbe, O.; Hamm, P. Kinetic response of a photoperturbed allosteric protein. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 11725−11730.

(57) Waldauer, S. A.; Stucki-Buchli, B.; Frey, L.; Hamm, P. Effect of viscogens on the kinetic response of a photoperturbed allosteric protein. *J. Chem. Phys.* **2014**, *141*, No. 22D514.

(58) Kumita, J. R.; Smart, O. S.; Woolley, G. A. Photo-control of helix content in a short peptide. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3803−3808.

(59) Hamm, P.; Kaindl, R. A.; Stenger, J. Noise suppression in femtosecond Mid-infrared light sources. *Opt. Lett.* **2000**, *25*, 1798−1800.

(60) Johnson, P. J. M.; Koziol, K. L.; Hamm, P. Intrinsic phasing of heterodyne-detected multidimensional infrared spectra. *Opt. Express* **2017**, *25*, 2928−2938.

(61) Zhang, J.; Sapienza, P. J.; Ke, H.; Chang, A.; Hengel, S. R.; Wang, H.; Phillipsand, G. N.; Lee, A. L. Crystallographic and nuclear magnetic resonance evaluation of the impact of peptide binding to the second PDZ domain of protein tyrosine phosphatase 1E. *Biochemistry* **2010**, *49*, 9280−9291.

(62) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindah, E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.

(63) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.

(64) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J. Comput. Chem.* **2000**, *21*, 132−146.

(65) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623−1641.

(66) Ernst, M.; Sittel, F.; Stock, G. Contact- and distance-based principal component analysis of protein dynamics. *J. Chem. Phys.* **2015**, *143*, No. 244114.

(67) Ernst, M.; Wolf, S.; Stock, G. Identification and validation of reaction coordinates describing protein functional motion: Hierarch-

ical dynamics of T4 lysozyme. *J. Chem. Theory Comput.* **2017**, *13*, 5076−5088.

(68) Lehner, F.; Kudlinzki, D.; Richter, C.; Müller-Werkmeister, H. M.; Eberl, K. B.; Bredenbeck, J.; Schwalbe, H.; Silvers, R. Impact of azidohomoalanine incorporation on protein structure and ligand binding. *ChemBioChem* **2017**, *18*, 2340−2350.

(69) Hell, S. W. Far-field optical nanoscopy. *Science* **2007**, *316*, 1153−1158.

DOI: 10.1021/acs.jpcb.8b08368
J. Phys. Chem. B 2018, 122, 10118−10125

Supplementary Materials: Azidohomoalanine: A Minimally-Invasive, Versatile and Sensitive Infrared Label in Proteins to Study Ligand Binding

Claudio Zanobini, Olga Bozovic, Brankica Jankovic, Klemens L. Koziol, Philip J. M. Johnson, Peter Hamm*
Department of Chemistry, University of Zurich, Zurich, Switzerland
and
Adnan Gulzar, Steffen Wolf, Gerhard Stock*
Biomolecular Dynamics, Institute of Physics, Albert Ludwigs University, Freiburg, Germany

**corresponding authors:* peter.hamm@chem.uzh.ch and stock@physik.uni-freiburg.de

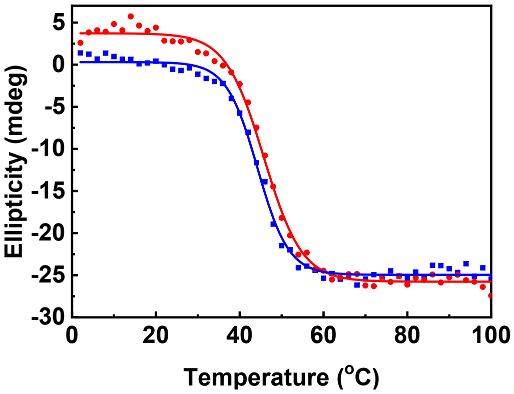


FIG. S1. CD signal wildtype PDZ2 (blue) and the K38Aha mutant (red) recorded at 205 nm. The solid lines show two-state fits to the data, revealing melting temperatures of 44 °C and 46 °C, respectively.

I. CIRCULAR DICHROISM SPECTROSCOPY

Circular dichroism spectra were collected using a Jasco J-810 spectropolarimeter in 1 mm quartz cuvette (Hellma Analytics, Quartz SUPRASIL) using 350 μL of protein solution. 20 μM protein solutions of both wild type and K38Aha PDZ2 domain were prepared in 50 mM borate, 150 mM NaCl, pH 8.5 buffer. The same buffer was used as a background. Temperature-dependent CD spectra were recorded at 205 nm and fitted in order to reveal a melting temperature of 46 ± 1°C (Fig. S1, red), somewhat larger than the corresponding value from the wild-type protein 44 ± 1°C (Fig. S1, blue).

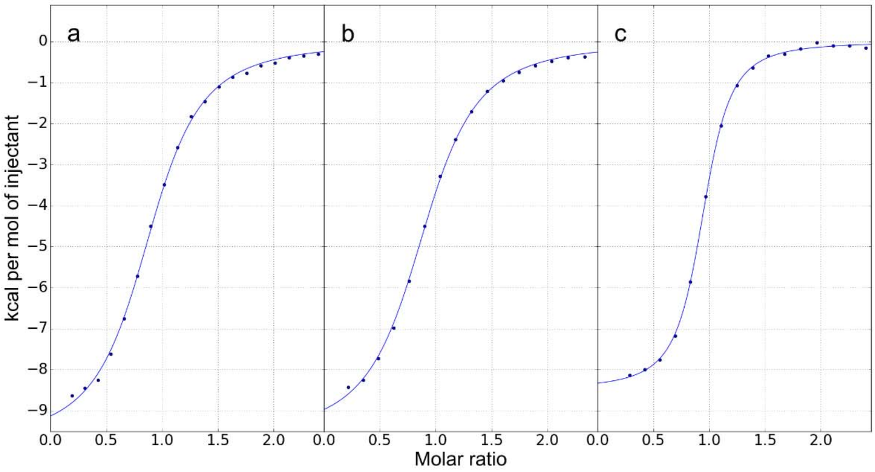


FIG. S2. ITC measurements of the binding affinity of the wild type peptide to (a) the wildtype PDZ2 domain and (b) its K38Aha mutant. Panel (c) shows the result for the binding affinity of the azo-peptide to the K38Aha mutant of PDZ2. In either case, the concentration of protein has been 140 μM. The data have been fitted as explained in the text, revealing binding constants of K_D=11μM for (a) and (b), and K_D=3μM for (c).

II. ISOTHERMAL TITRATION CALORIMETRY

Isothermal titration calorimetry was performed using MicroCal^{TM} iTC200 system (Malvern Instruments Ltd, Malvern, UK) at 21 °C (Fig. S2). Immediately before the experiment was performed, protein and peptide solutions were extensively re-dialyzed against 50 mM borate, 150 mM NaCl, pH 8.5 buffer. The sample cell of the calorimeter was loaded with 250 μL of 0.14 mM protein solution in 50 mM borate, 150 mM NaCl, pH 8.5 buffer. Circa 40 μL of 1.4 mM peptide solution in the same buffer was added in small increments (1.5 μL) with 120 seconds spacing between injections until the whole sigmoidal thermogram curve was obtained (26 injections). Stirring speed was set to 1000 rpm. Heats of dilution were determined by titrating the same peptide solution into the dialysis buffer and subtracted from the raw titration data before data analysis. The data were processed using the Origin ITC software provided by the instrument manufacturer. Fitting was performed using one set of sites binding model and K_D value was obtained directly from the fitted binding curve.

III. INDEPENDENT DATA SET

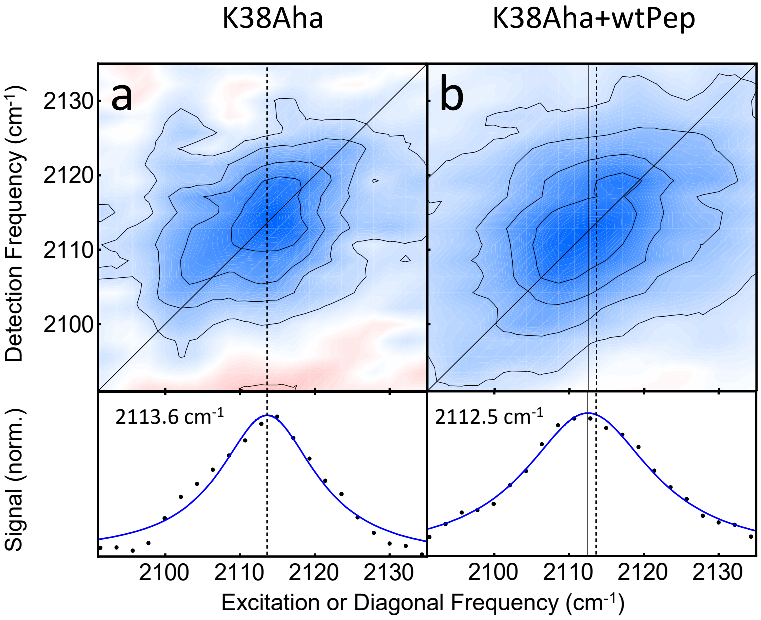


FIG. S3. (a) 2D IR response of K38Aha and (b) K38Aha+wtPep. In this case, the protein concentration has been 0.5 mM and the peptide concentration 2.5 mM. In panel (a) the dotted line marks the reference peak position of of K38Aha and in panel (b) the peak positions of K38Aha+wtPep is marked in addition as solid lines. These data are, in principle, the same as in Fig. 2ab, but has been measured at a different day by different people. The deduced peak frequencies (2113.6 cm^{-1} and 2112.5 cm^{-1}, respectively) are the same within 0.1 cm^{-1} as in Fig. 2ab, evidencing the reproducibility of these measurements.

IV. MD CONTACT ANALYSIS

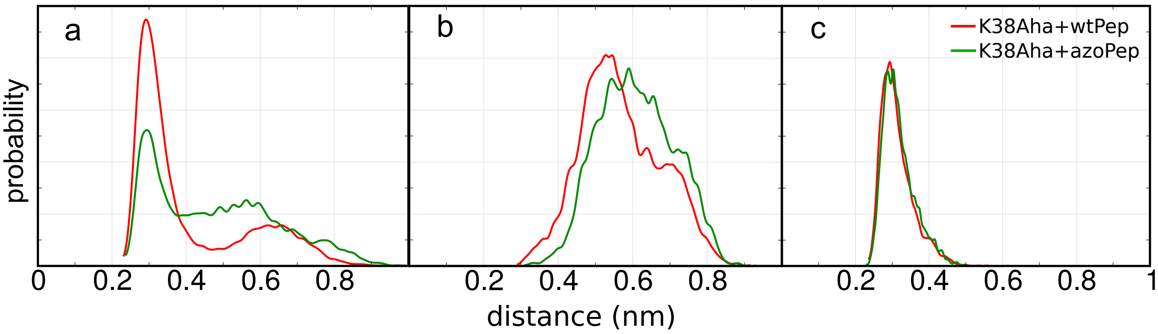


FIG. S4. Probability distributions of minimum distances between ligand residues (a) Glu(-5), (b) Gln(-4) and (c) Val(-3) and the azido group of Aha38.

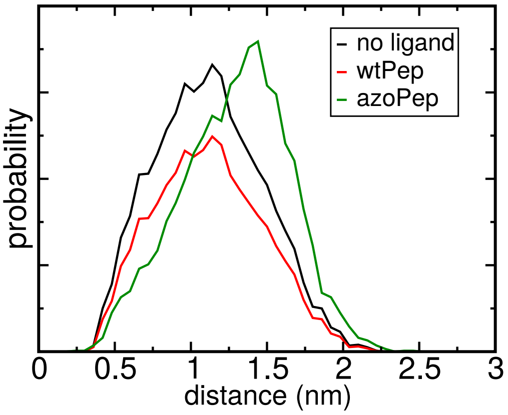


FIG. S5. Probability distributions of minimum distances between the azido group of Aha38 and any ion in the simulation box.

V. AZO SWITCH PARAMETERS

Azo-switch parameters were obtained using Antechamber[1] and Acpype.[2] For parameterization, a model of the switch with attached cysteine side chains was built in Pymol[3] based on a preexisting structure with additional sulfonate side chains.[4] Atomic parameters were derived from GAFF parameters[5] with BCC charges,[6,7] the latter of which have been found to be consistent with the Amber protein force field used in the simulation.[8–10]

[AZO]

[atoms]

N	N	-0.41570	1
H	H	0.27190	2
CA	CT	0.02130	3
HA	H1	0.11240	4
CB	CT	-0.12310	5
HB1	H1	0.11120	6
HB2	H1	0.11120	7
SG	S	-0.05310	8
CD	CT	-0.32850	9
HD1	H1	0.15125	10
HD2	H1	0.15125	11
CE	C	0.30800	12
OE	O	-0.34000	13
NZ	N	-0.31550	14
HZ	H	0.24550	15
CH	CA	0.09400	16
CT1	CA	-0.18075	17
HT1	HA	0.15450	18
CI2	CA	-0.06850	19
HI2	HA	0.15100	20
CT2	CA	-0.18075	21
HT2	HA	0.15450	20
CI1	CA	-0.06850	26
HI1	HA	0.15100	27
CK1	CA	-0.07400	28
NL	NE	-0.07000	29
C	C	0.59730	30
O	O	-0.56790	31

```
[ bonds ]

        N       H
        N       CA
        CA      HA
        CA      CB
        CA      C
        CB      HB1
        CB      HB2
        CB      SG
        SG      CD
        CD      HD1
        CD      HD2
        CD      CE
        CE      OE
        CE      NZ
        NZ      HZ
        NZ      CH
        CH      CT1
        CH      CT2
        CT1     HT1
        CT1     CI1
        CI1     HI1
        CI1     CK1
        CT2     HT2
        CT2     CI2
        CI2     HI2
        CI2     CK1
        CK1     NL
        C       O
        -C      N

[ impropers ]

    -C      CA      N       H
    CA      +N      C       O
    N       CA      C       +N    105.4    0.75    1

[ atomtypes ]

      NE    7    14.01    0.0000    A    3.25000e-01    7.11280e-01

[ bondtypes ]

     NE    NE    1     0.12640    604239.0 ; AZO
     NE    CA    1     0.14310    302956.8 ; AZO
     CA    N     1     0.14220    311749.1 ; AZO
```

```
    [ angletypes ]

    NE     NE     CA     1        115.190     583.305 ; For the AZO NE NE CA
    NE     CA     CA     1        119.880     567.228 ; For the AZO NE CA CA
    CT     CA     CB     1        119.450     542.330 ; CB - CG - CD2 AZU GAFF
    CA     CB     CA     1        114.190     570.360 ; CG - CD2 - CE2 AZU GAFF

    [ dihedraltypes ]

CA      CA     N      H      9        180.0       1.88406    2 ; AZO side chain connection
NE      NE     CA     CA     9        180.0       0.00000    3 ; AZO side chain connection
CA      CA     N      C      9        180.0       1.88406    2 ; AZO side chain connection
```

References:

---

[1] Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations, *J. Mol. Graph. Model.* **2006**, *25*, 247–260.

[2] Sousa da Silva, A. W.; Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE, *BMC Res. Notes* **2012**, *5*, 367.

[3] Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8 **2010**.

[4] Stucki-Buchli, B.; Johnson, P. J. M.; Bozovic, O.; Zanobini, C.; Koziol, K. L.; Hamm, P.; Gulzar, A.; Wolf, S.; Buchenberg, S.; Stock, G. 2D-IR spectroscopy of an AHA labelled photoswitchable PDZ2 domain, *J. Phys. Chem. A* **2017**, *121*, 9435–9445.

[5] Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field, *J. Comput. Chem.* **2004**, *25*, 1157–1174.

[6] Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method, *J. Comput. Chem.* **2000**, *21*, 132–146.

[7] Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, *J. Comput. Chem.* **2002**, *23*, 1623–1641.

[8] Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, *Expert Opin. Drug Dis.* **2015**, *10*, 449–461.

[9] Segala, E.; Guo, D.; Cheng, R. K. Y.; Bortolato, A.; Deflorian, F.; Doré, A. S.; Errey, J. C.; Heitman, L. H.; IJzerman, A. P.; Marshall, F. H.; Cooke, R. M. Controlling the Dissociation of Ligands from the Adenosine A2A Receptor through Modulation of Salt Bridge Strength., *J. Med. Chem.* **2016**, *59*, 6470–6479.

[10] Ruiz-Carmona, S.; Schmidtke, P.; Luque, F. J.; Baker, L.; Matassova, N.; Davis, B.; Roughley, S.; Murray, J.; Hubbard, R.; Barril, X. Dynamic undocking and the quasi-bound state as tools for drug discovery., *Nature Chem.* **2017**, *9*, 201–206.

Chapter 4

Real-Time Observation of Ligand-Induced Allosteric Transitions in a PDZ Domain

After exploring the PDZ2 domain recognition of two diverse peptide-ligands in a minimally invasive manner by combining azidohomoalanine (Aha) site-specific infrared labeling with 2D-IR spectroscopy [35], the photoswitchable peptide-ligand version was employed to promote a non-equilibrium allosteric transition of the PDZ2 domain. Both *cis* to *trans* and *trans* to *cis* photoswitching approaches were investigated, observing the amide I response of the protein-peptide adduct by transient infrared (TRIR) spectroscopy. An isotopically labeled (^{13}C^{15}N) PDZ2 domain variant was expressed, allowing for the isolation of the PDZ2 domain response from the one of the photoswitchable ligand, as a result of the red-shifted amide I response of the heavy (^{13}C^{15}N) isotopologue respect to the one of the naturally abundant (^{12}C^{14}N) PDZ2 domain. The photoswitchable ligand has a different binding affinity towards the PDZ2 domain of *circa* 5 fold regarding the *trans* or *cis* isomerization state of its azobenzene moiety, where the binding affinity is higher (smaller K_D) in the *trans* configuration. A change of the binding affinity is a prerequisite to expect a protein structural rearrangement together with a change of the side-chain fluctuation dynamics, which are related to the allosteric signal propagation in PDZ domains [59–65], as it happens in nature by phosphorylation of the ligand or upon ligand binding versus unbinding. The structural rearrangements of the protein induced by ligand photoswitching were acquired until ≈ 42 μs, revealing asymmetric non-equilibrium evolution pathways for the two photoswitching directions. Moreover from the lifetime analysis, a small number of discrete time scales emerges for each photoswitching approach, highlighting in this way the time scale separation involved in the overall non-equilibrium allosteric transition of the PDZ2 domain. These experimental results can be interpreted with a Markov state model [183–188], where the PDZ2 domain structural ensemble can be populated differently depending on the *cis* to *trans* or *trans* to *cis* photoswitching approach.

∗ The manuscript "Real-Time Observation of Ligand-Induced Allosteric Transitions in a PDZ Domain" is in preparation. The authors of the manuscript are: Olga Bozovic[1,†], Claudio Zanobini[1,†], Adnan Gulzar[2,†], Brankica Jankovic[1], David Buhrke[1], Matthias Post[2], Steffen Wolf[2], Gerhard Stock[2,∗], Peter Hamm[1,∗]
[1]*Department of Chemistry, University of Zurich, Zurich, Switzerland*
[2]*Institute of Physics, Albert Ludwigs University, Freiburg, Germany*
[†]*authors contributed equally*
[∗]*corresponding authors: stock@physik.uni-freiburg.de, peter.hamm@chem.uzh.ch*
I contributed to the work present in this chapter with the acquisition, elaboration, and analysis of the transient infrared data in Fig. 4.4, 4.5, 4.6, and with the paper writing.

4.1 Introduction

Allosteric regulation is a fundamental mechanism that nature employs to control protein activity; it implies *e.g.* the binding of a small molecule to a protein which acts as a regulatory trigger for the latter. The binding event takes place in a zone of the protein (the allosteric site) distant from the region where the effects of it manifest (active site). In recent years the concept of allostery became wider, including not only its initial definition where the allosteric interaction provokes a conformational change of the protein but also in terms of structural fluctuations as proposed by *Cooper* and *Dryden* [189]. The allosteric and active sites can be several ångströms (Å) far away from each other, and an inevitable question arises: how does the allosteric signal propagate through the protein? It would be interesting to know the answer to this fundamental scientific question.

NMR spectroscopy is a useful tool to provide us with information about protein structures [190] and structural fluctuations at the equilibrium, employing different relaxation methods of magnetically active nuclei [66, 67]. Due to its intrinsic time resolution limitations, the non-equilibrium dynamics of proteins have not been investigated extensively yet, especially for what it concerns the phenomenon of allostery [145, 191]. On the other hand, TRIR spectroscopy is a useful technique which can be employed to study the non-equilibrium dynamics of proteins [145], due to its inherent high time resolution.

PDZ domain are protein-protein interaction modules which are involved in signaling transduction complexes and for this reason they are deeply regulated [57–59, 76–85]. These modules show a modest conformational change upon ligand binding ($\lesssim 1$ Å) [57, 91], and represent a striking example where changes of the equilibrium structural fluctuations are involved in the allosteric process [59–65]. Among the various regulation mechanisms of proteins like phosphorylation, disulfide bond formation/breaking, change in protein flexibility, *etc.* [192], ligand phosphorylation plays a pivotal role in the regulation of PDZ domains [60, 193, 194]. It has been reported that ligand phosphorylation in the case of Tiam1 PDZ domain/Phospho-Syndecan1 leads to a damping effect of the dynamic motions of the distal protein regions, despite the small change in binding affinity after phosphorylation of the ligand (K_D folds difference < 2) [60]. In the case of the PDZ2 domain, which is the object of study and discussion of this chapter, phosphorylation of the ligand provokes $\approx$ 5-7 fold difference in binding affinity [193]. It has to be underlined that, a change in binding affinity is often a prerequisite to induce an allosteric effect in PDZ domains [60, 194].

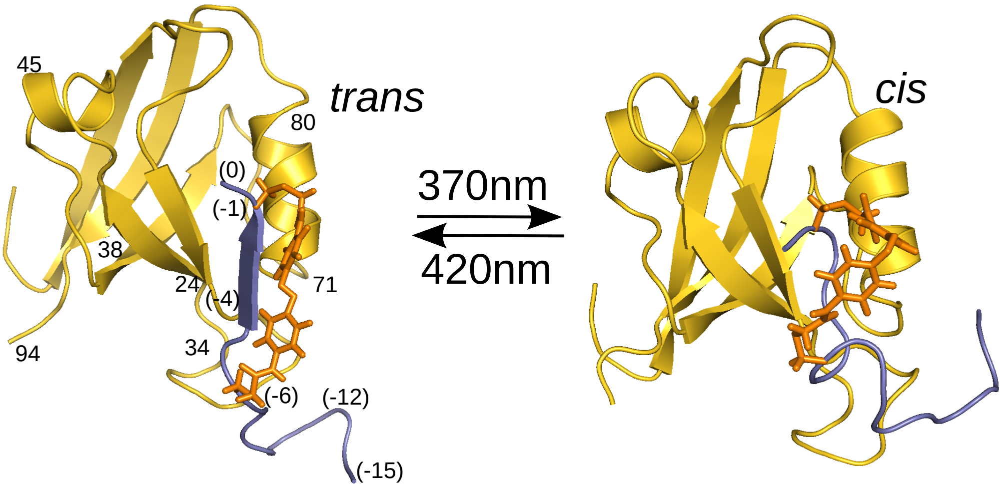


Figure 4.1: Wild type PDZ2 domain with the photoswitchable peptide bound to it in both *trans* and *cis* configurations. Upon illumination at two different wavelengths (370 and 420 nm), the *trans* to *cis* and the *cis* to *trans* photoisomerizations can be promoted.

Here, a novel photocontrollable system is presented: it consists of the PDZ2 domain from hPTP1E (human tyrosine phosphatase 1E) and a photoswitchable RA-GEF-2 peptide version (Ras/Rap1 associating guanidine nucleotide exchange factor 2). The photoisomerization of an azobenzene moiety, which is present in the structure of the photoswitchable peptide-ligand, can be promoted in both *trans* to *cis* and *cis* to *trans* directions reversibly, employing two different wavelengths at 370 nm and 420 nm, respectively (see Fig. 4.1). The photoswitch used in this case to be covalently linked to the peptide sequence RWAKSEAKECEQVSCV is the water-insoluble version without the sulfonate groups ($—SO_3^-$) [143] reported on the top of Fig. 1.6. The water-insoluble photoswitch does not compromise the water-solubility of the linking product at mM concentrations, values needed to perform transient infrared experiments, due to the high hydrophilicity of this peptide. Moreover, the binding affinity of the photoswitchable peptide-ligand towards the PDZ2 domain shows a *circa* 5 fold difference between the *trans* and *cis* configurations of the photoswitch, which mimics to almost the same extent what it happens in nature *via* phosphorylation [193]. The dissociation constants (K_D) have been determined employing three different methods: isothermal titration calorimetry (ITC), intrinsic tryptophan (W) fluorescence quenching and circular dichroism (CD) spectroscopy (see Fig. 4.2, 4.3, and Methods for details – section A.1 – Determination of the Binding Affinities) [148]. The K_D and their mean value obtained from these methods ($K_{D,trans} = 2.0 \pm 0.6$ μM, $K_{D,cis} = 9.6 \pm 0.5$ μM) have been reported in Tab. 4.1. It has to be stressed that the photoswitch element, which is needed to promote the non-equilibrium perturbation of the PDZ2 domain, is covalently linked to the peptide and not to the protein, leaving the latter intact without introducing any artificial element on it. Furthermore, to unravel the highly convoluted amide I response of the PDZ2 domain from the one of the photoswitchable peptide, isotope labeling strategy has been employed [46–50], expressing a (^{13}C^{15}N) PDZ2 domain variant. Acting in this way, the ultrafast transient amide I response of the PDZ2 domain upon peptide photoswitching has been isolated, obtaining an experimental observable which can be correlated with the allosteric signal propagation in the PDZ2 domain.

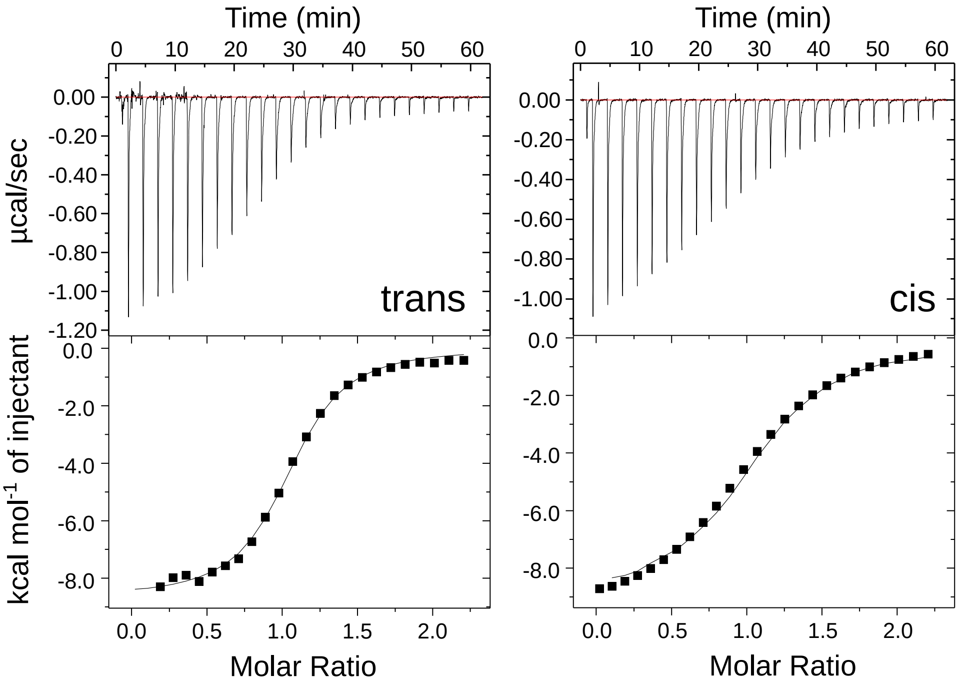


Figure 4.2: Isothermal titration calorimetry (ITC) thermograms of the wild type PDZ2 domain with the photoswitchable peptide-ligand in the *trans* and *cis* configurations (left and right plot, respectively).

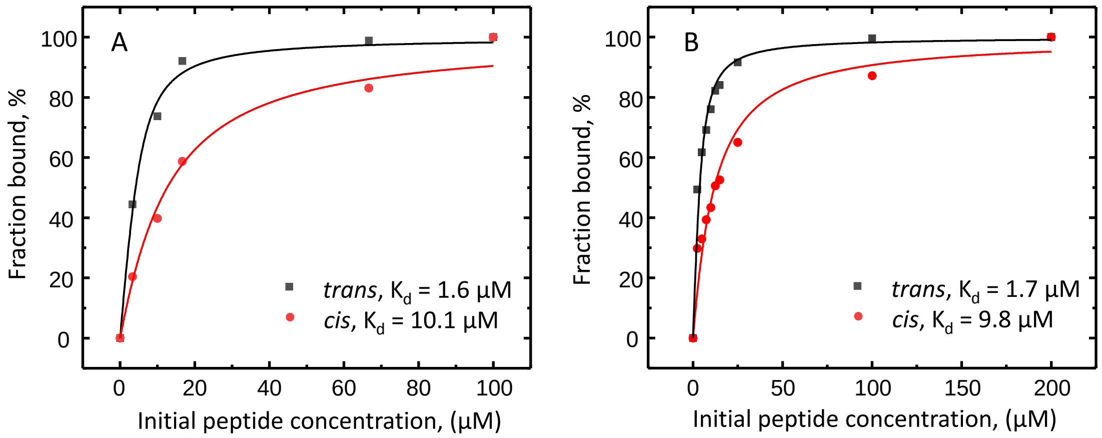


Figure 4.3: Dissociation constants (K_D) determination for the wild type PDZ2 domain and the photoswitchable peptide-ligand in the *cis* (red) and *trans* (black) configurations, employing circular dichroism (CD) spectroscopy and intrinsic tryptophan (W) fluorescence quenching methods (panels A and B, respectively).

Table 4.1: Dissociation constants (K_D) for the wild type PDZ2 domain and the photoswitchable peptide-ligand in the *cis* and *trans* configurations obtained from isothermal titration calorimetry (ITC), circular dichroism (CD) spectroscopy and intrinsic tryptophan (W) fluorescence quenching methods. The K_D and their mean value have been reported in the following table.

	K_D (μM)	
	trans	*cis*
ITC	2.7 ± 0.2	9.1 ± 0.4
CD	1.6 ± 0.1	10.1 ± 0.3
Fluorescence	1.7 ± 0.3	9.8 ± 1.2
Mean	2.0 ± 0.6	9.6 ± 0.5

4.2 Pump-Probe Measurements and Lifetime Analysis

The system under investigation is shown in Fig. 4.1. This figure is representative of the experiment performed, where two main considerations had to be taken into account:

- Due to the high concentrations needed to perform the transient infrared experiments (photoswitchable peptide ≈ 1.25 mM, PDZ2 domain ≈ 1.5 mM) together with the K_D values of ≈ 2.0 μM and ≈ 9.6 μM for the *trans* and *cis* states, the photoswitchable peptide is essentially bound to the PDZ2 domain in both *trans* and *cis* configurations (fraction bound > 95% in both cases).

- In these conditions, the difference between the *trans* and *cis* states in terms of fraction bound of the photoswitchable peptide to the PDZ2 domain is negligible ($< 5\%$). Moreover, the timescales for unbinding events are most probably beyond the time window accessible with our experiment (*e.g.* $k_{off} \approx$ 40-70 s^{-1} for RA-GEF-2 peptide) [195].

The pump-probe spectra (see Methods for details – section A.2 – Transient Infrared (IR) Spectroscopy) were collected in the amide I spectral region for both *trans* to *cis* and *cis* to *trans* photoswitching directions; the data are shown in the panels (d-f) and (g-i) of Fig. 4.4, respectively. The amide I vibration at $\approx$ 1650 cm^{-1} is known to be sensitive to the secondary structure of proteins and peptides [45]. It originates mainly from the C=O stretching vibration, depending on the relative orientation and distances among the C=O dipoles in the backbone structure. A change of the secondary structure affects the resonance interaction among the oscillating C=O dipoles, which causes a change of the amide I absorption band [45].

The transient infrared spectra reported in Fig. 4.4 were normalized to each other based on several parameters: the different pump-pulse energies employed for the *trans* to *cis* and *cis* to *trans* photoswitching experiments (*circa* 2.1 and 1.3 μJ, respectively), the characteristic extinction coefficient values of the *trans* and *cis* photoswitchable peptide isomers (*circa* 23500 M^{-1} cm^{-1} for *trans* at 380 nm versus 2000 M^{-1} cm^{-1} for *cis* at 420 nm) [131], and the distinct isomerization quantum yields of each photoswitching direction (*circa* 8% for *trans* to *cis* and 62% for *cis* to *trans*, respectively) [141]. This kind of normalization means that the transient infrared data shown in Fig. 4.4 refers to the same amount of isomerized molecules in both photoswitching experiments. The left panels (a, d, g) of Fig. 4.4 show the results for the wild type PDZ2 domain bound to the photoswitchable peptide and the middle panels (b, e, h) those for the analogous sample with the same concentrations of the two species as before but with an isotopically labeled ($^{13}C^{15}N$) PDZ2 domain variant, which red-shifts the amide I absorption band of the protein upon peptide photoswitching. Even though the transient infrared responses in panels (d, e) and (g, h) for the two photoswitching approaches appear similar, they are not, and the desired information is hidden within them. The observed signals are dominated by the peptide response, which is covalently linked to the azobenzene moiety of the photoswitch, the element responsible for the photoisomerization process. In order to isolate the smaller PDZ2 domain response upon peptide photoswitching from the dominant response of the photoswitchable peptide, the transient data have been subtracted, exploiting the fact that only the amide I response of the PDZ2 domain has been red-shifted due to the isotope labeling effect [46–50], and obtaining the data in the panels (c, f, i) of Fig. 4.4. It has to be stressed that the sample concentrations have been kept the same in the various transient infrared experiments

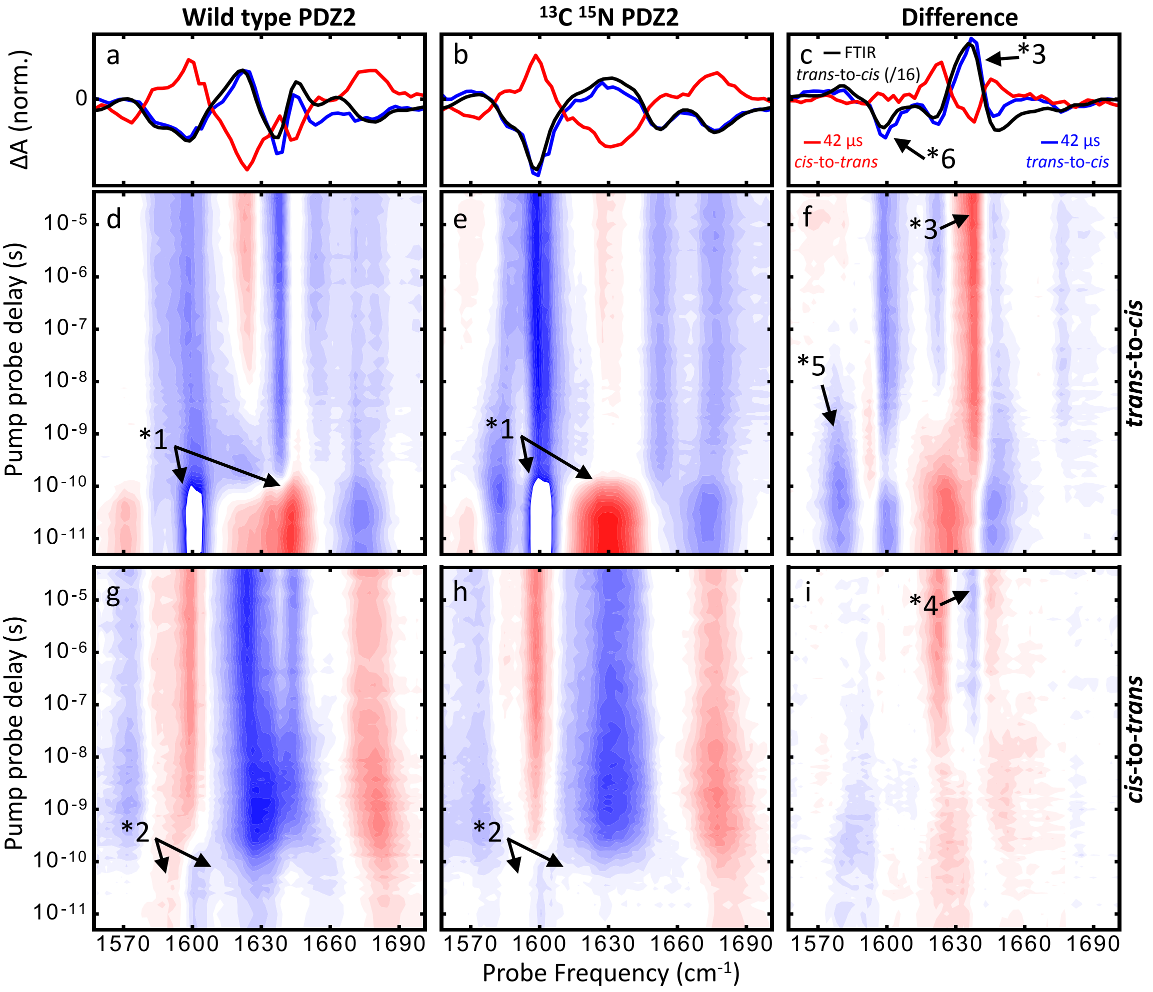


Figure 4.4: Transient infrared spectra in the amide I spectral region. Panels (a-c) compare the transient data acquired at long pump-probe delay times (averaged from $\approx$ 20 to 42 μs to improve the signal to noise) for *trans* to *cis* (blue line) and *cis* to *trans* (red line) photoswitching, together with the properly scaled *trans* to *cis* FTIR difference spectrum (black line). Panels (d-f) and (g-i) show the transient data for *trans* to *cis* and *cis* to *trans* photoswitching, respectively. Left panels (a, d, g) show the transient data for the sample with the wild type (WT) protein, middle panels (b, e, h) for the sample with the protein ^{13}C^{15}N-labeled (the photoswitchable peptide is with naturally abundant ^{12}C^{14}N), and right panels (c, f, i) for the ^{13}C^{15}N-WT difference spectra. Red colours in panels (d-i) indicate positive absorbance changes, blue colours negative absorbance changes. The relative scaling of the two data sets for the two photoswitching directions as well as the labeled features are described in the text.

and that the experiments themselves have been performed right after each other for each photoswitching direction, without changing any experimental parameter.

The pump-probe data shown in the panels (d-f) of Fig. 4.4 for the *trans* to *cis* experiment manifest strong signals at early delay times up to *circa* 100 ps (marked as *1 and some of them have been highlighted in white) which are the consequence of the vibrational cooling of the photoswitch after the absorption of the excita-

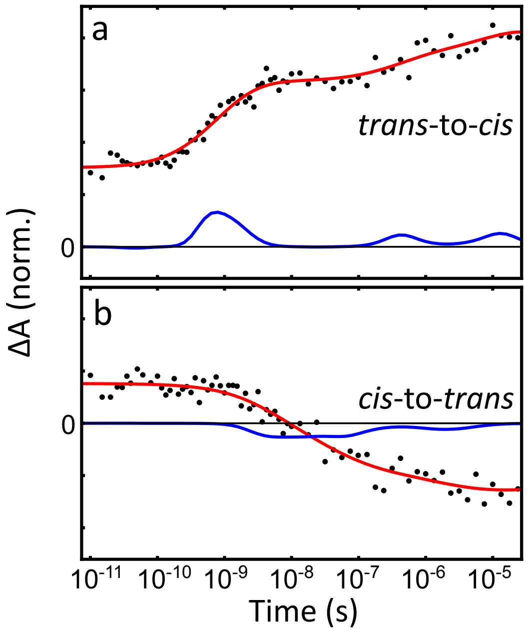


Figure 4.5: Transient kinetic traces of the $^{13}C^{15}N$-WT difference data at ≈ 1636 cm^{-1}, *i.e.* they represent the isolated transient infrared responses of the band marked with *3 and *4 in Fig. 4.4 of the $^{13}C^{15}N$-labeled PDZ2 domain upon peptide photoswitching. Panel (a) and (b) show the results for the *trans* to *cis* and the *cis* to *trans* photoswitching directions, respectively. The experimental data are shown as black dots, the fit $S(\omega_i, t)$ (see Methods for details, section A.3 – Lifetime Analysis – Eq. A.1) as red lines, and the amplitudes $a(\omega_i, \tau_j)$ of the lifetime spectra as blue lines.

tion pulses needed to promote the photoisomerization event [25]. In other words, the azobenzene moiety, which constitutes the photoswitch, dissipates part of the absorbed energy to its chemical environment in the form of heat with a different extent depending on the quantum yields for the *trans* to *cis* and *cis* to *trans* photoswitching processes. This kind of effect is known, has been already observed before [25, 72, 136] and have a relaxation time of *circa* 10 ps. This time scale is faster than the pump-pulse duration which is *circa* 200 ps (the pump-pulse was stretched to this value bypassing the compression stage in the regenerative amplifier to minimize the sample degradation during the transient infrared experiments), which signifies that the time response of the heat signal reflects the temporal shape of the pump-pulse. The heat signal is bigger for the *trans* to *cis* process due to the smaller quantum yield in this case respect to the *cis* to *trans* photoisomerization, which *de facto* means that a smaller fraction of absorbed photons are productive for the *trans* to *cis* process [72] and that the heat signal is significantly smaller in the *cis* to *trans* case (see marks *2 in panels (g, h) of Fig. 4.4).

The isolated PDZ2 domain non-equilibrium responses upon *trans* to *cis* and *cis* to *trans* peptide photoswitching (see panels (f, i) in Fig. 4.4) show complex spectral features and are spread over several orders of magnitude in time. Moreover, the

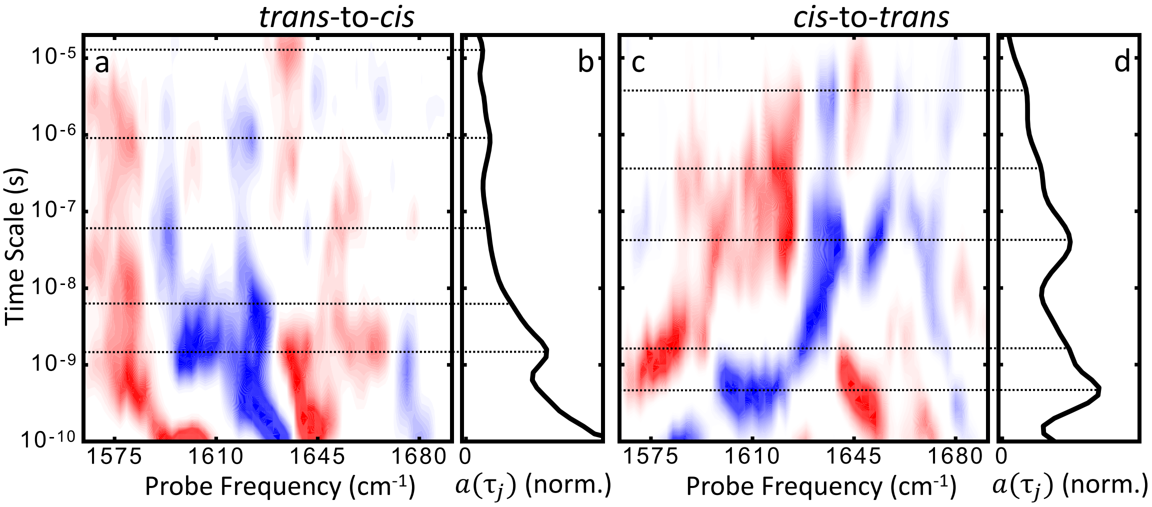


Figure 4.6: Lifetime analysis for the ^{13}C^{15}N-WT difference data in Fig. 4.4. Panels (a, c) show the lifetime spectra $a(\omega_i, \tau_j)$ (see Methods for details, section A.3 – Lifetime Analysis – Eq. A.1) of the ^{13}C^{15}N-WT difference data for *trans* to *cis* and *cis* to *trans* photoswitching, respectively. Red colours indicate positive amplitudes $a(\omega_i, \tau_j)$, while blue colours negative amplitudes $a(\omega_i, \tau_j)$. Panels (b, d) show the averaged lifetime spectra $a(\tau_j)$ (black thick lines) for *trans* to *cis* and *cis* to *trans* photoswitching, respectively. The 2D information present in panels (a, c) has been averaged in an absolute sense over the probe frequency axis ω_i and squeezed in a 1D plot (see Eq. A.5). The horizontal black dotted lines connecting the lifetime spectra $a(\omega_i, \tau_j)$ in panels (a, c) with the corresponding average lifetime spectra $a(\tau_j)$ in panels (b, d) highlight distinct timescales for the two photoswitching directions.

two responses are not mirror-images of each other, which would be expectable if the response of the PDZ2 domain evolved through the same pathway upon *trans* to *cis* and *cis* to *trans* photoisomerizations. There are some examples in the literature which suggest that different evolution pathways, for the forward and backward reactions of a determined process, are characteristic of non-equilibrium phenomena [72, 196–199]. For instance, the most pronounced band at $\approx$ 1636 cm^{-1} (marked as *3) appears at early pump-probe delay times (at 1 ns it is already present) in the *trans* to *cis* experiment, while the same band (marked as *4) develops at $\approx$ 10-100 ns in the *cis* to *trans* photoswitching direction. The transient kinetic traces of the band mentioned above are shown in Fig. 4.5 and from their comparison, the distinct dynamics highlight the different pathways followed by the PDZ2 domain response with the two photoswitching approaches. Furthermore, the transient band at $\approx$ 1580 cm^{-1} (marked as *5) in panel (f) for the *trans* to *cis* experiment has a lifetime of $\approx$ 10 ns, which has no counterpart in panel (i) for the *cis* to *trans* acquisition, namely an equivalent transient band with opposite sign.

The data in the panels (a-c) in Fig. 4.4 show the comparison between the transient infrared data acquired at long pump-probe delay times (averaged from $\approx$ 20 to 42 μs to improve the signal to noise) for the *trans* to *cis* (blue line) and *cis* to *trans* (red line) experiments, together with the properly scaled *trans* to *cis* FTIR difference spectrum (black line, FTIR divided by 16), which represents *de facto* the response at infinite time after photoswitching. The aforementioned transient infrared data have similar intensities (after normalizing them for the same number of photoswitched molecules as described before) without being exact mirror-images of each other. Above all from their comparison, it sticks out the negative band in blue (marked as *6) at ≈ 1600 cm^{-1} for the *trans* to *cis* spectrum, which counterpart in the red *cis* to *trans* spectrum is absent. It is remarkable to note that the black FTIR difference spectrum reveals this negative band at ≈ 1600 cm^{-1}. These experimental observations allows us to conclude that in both photoswitching experiments, the PDZ2 domain undergoes significant structural rearrangements within 42 μs, namely the non-equilibrium response of the protein is confined mostly within this time scale. The tighter match of the blue *trans* to *cis* spectrum to the black FTIR difference spectrum in panel (c) of Fig. 4.4 indicates that the progress of the overall non-equilibrium process after 42 μs is more forward for the *trans* to *cis* than for the *cis* to *trans* photoswitching experiment.

In order to extract the maximum information from the experimental data, the maximum entropy inversion of the Laplace transform [200, 201] has been employed to analyse the isolated PDZ2 domain responses upon *trans* to *cis* and *cis* to *trans* peptide photoswitching reported in panels (f, i) of Fig. 4.4, respectively. This analysis gave rise to the lifetime spectra in panels (a, c) and to the corresponding averaged lifetime spectra in panels (b, d) of Fig. 4.6 for the two photoswitching approaches (see Methods for details, section A.3 – Lifetime Analysis). The lifetime analysis represents the dynamic evolution of the non-equilibrium kinetics of the isolated protein responses, which is related to the allosteric signal propagation in the PDZ2 domain, as a consequence of the perturbations induced by peptide photoswitching in the *trans* to *cis* and *cis* to *trans* directions. From the lifetime analysis in Fig. 4.6, two observations are worth to be mentioned:

1. The non-equilibrium response of the protein is asymmetric; there are no regular counterparts between the amplitudes of the lifetime spectra in panels (a, c) for the two photoswitching experiments.

2. A discrete set of timescales (marked as horizontal black dotted lines) emerges for each *trans* to *cis* and *cis* to *trans* photoswitching approach, where both sets are associated to the maxima present in each of the corresponding average lifetime spectra in panels (b, d). These maxima represent the time scales to which the transient structural changes of the PDZ2 domain are the biggest.

From the second remark, a discrete set of timescales is characteristic of each photoswitching direction, which can be interpreted with a Markov state model representation of the overall process [183–188], where the horizontal black dotted lines in Fig. 4.6 represent the most populated structures of the PDZ2 domain structural ensemble for each *trans* to *cis* and *cis* to *trans* photoswitching approach. Molecular dynamics (MD) simulations are underway to correlate the non-equilibrium protein response with the allosteric signal propagation in the PDZ2 domain.

Chapter 5

Photocontrolling Protein-Peptide Interactions: from Minimal Perturbation to Complete Unbinding

In parallel with the studies on the PDZ2 domain, another system was investigated; the RNase S complex [74, 75, 175, 176, 179, 202]. Five photoswitchable S-peptide variants were designed and synthesized with the intent to modulate the binding affinity between the two partners which constitute the RNase S adduct, maximizing the dissociation constant (K_D) difference upon *cis* to *trans* photoisomerization. The foremost design criteria for the realization of such photoswitchable S-peptides was the decrease of the amount of α-helical content upon *cis* to *trans* photoisomerization [143]. From simulation studies of our collaborators in Freiburg (Prof. Dr. Gerhard Stock and coworkers), the increase of the experimentally observed K_D values upon *cis* to *trans* photoswitching correlates well with the reduction of α-helicity of the various photoswitchable S-peptide versions. Regarding the S-pep(6,13) variant, non-specific binding was observed for one of the two photoswitch states (the *trans* configuration) [148, 203, 204], revealing an ideal candidate for unbinding time-resolved experiments. The transient infrared data have been reported in a separate section of this chapter.

$\ast$ I contributed to this paper with the design of the photocontrollable system RNase S concerning S-pep(6,15), the purification of S-protein, the troubleshooting regarding the K_D determination of the various S-peptide mutants, the realization of a highly sensitive optical fluorescence setup for the K_D determination of S-pep(6,13), and with Fig. S1 and S2.

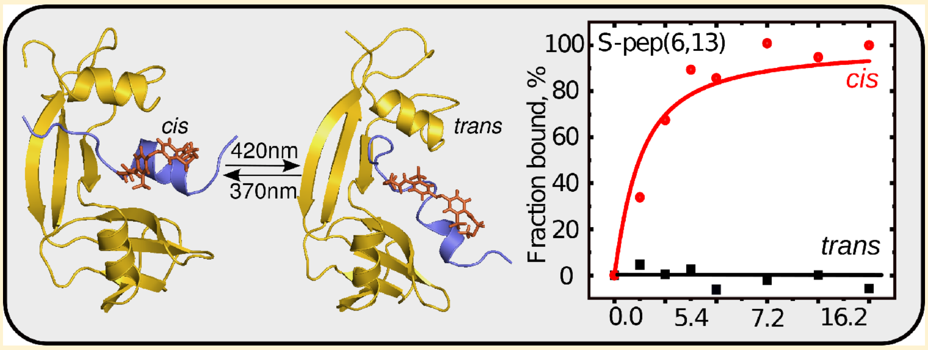
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
Cite This: J. Am. Chem. Soc. 2019, 141, 10702−10710
pubs.acs.org/JACS

Photocontrolling Protein−Peptide Interactions: From Minimal Perturbation to Complete Unbinding

Brankica Jankovic,[†,§] Adnan Gulzar,[‡,§] Claudio Zanobini,[†] Olga Bozovic,[†] Steffen Wolf,[‡] Gerhard Stock,*,[‡] and Peter Hamm*,[†]

[†]Department of Chemistry, University of Zurich, Zurich CH-8057, Switzerland
[‡]Biomolecular Dynamics, Institute of Physics, Albert Ludwigs University, Freiburg 79104, Germany

Supporting Information



ABSTRACT: An azobenzene-derived photoswitch has been covalently cross-linked to two sites of the S-peptide in the RNase S complex in a manner that the α-helical content of the S-peptide reduces upon *cis*-to-*trans* isomerization of the photoswitch. Three complementary experimental techniques have been employed, isothermal titration calorimetry, circular dichroism spectroscopy and intrinsic tyrosine fluorescence quenching, to determine the binding affinity of the S-peptide to the S-protein in the two states of the photoswitch. Five mutants with the photoswitch attached to different sites of the S-peptide have been explored, with the goal to maximize the change in binding affinity upon photoswitching, and to identify the mechanisms that determine the binding affinity. With regard to the first goal, one mutant has been identified, which binds with reasonable affinity in the one state of the photoswitch, while specific binding is completely switched off in the other state. With regard to the second goal, accompanying molecular dynamics simulations combined with a quantitative structure activity relationship revealed that the α-helicity of the S-peptide in the binding pocket correlates surprisingly well with measured dissociation constants. Moreover, the simulations show that both configurations of all S-peptides exhibit quite well-defined structures, even in apparently disordered states.

1. INTRODUCTION

Understanding conformational and dynamic aspects of protein−protein interactions is of paramount importance for new strategies in controlling them. In order to employ time-resolved techniques for the investigation of conformational changes of proteins and/or peptides, e.g., by transient X-ray scattering[1] or transient infrared (IR) spectroscopy,[2] fast triggering of a given process is required. This is where the importance for external control of specific interactions emerges: it facilitates detailed understanding of real-time conformational changes induced by protein−peptide recognition. Photocontrol has been recognized as a valuable tool for the external manipulation of numerous *in vitro* as well as *in vivo* processes. This way of triggering conformational changes offers numerous advantages, in particular high spatial and temporal resolution and selectivity.[3] Photoswitching, as one of the most promising approaches of photocontrol, has the additional advantage of being reversible.[4] For example, a reversible photocontrol of the α-helical content of small peptides with sufficient α-helical propensity can be achieved by cross-linking to sites of the peptide with an azobenzene-based molecule.[5−8] Upon illumination, the azobenzene chromophore undergoes a reversible *cis* to *trans* isomerization around its central diazene (N$=$N) double bond. Isomerization changes the geometry of the azobenzene moiety with different end-to-end distances of the two configurations, leading to either a perturbation or a stabilization of the helix, depending on the distance between sites to which the azobenzene moiety is bound.[9] Besides triggering folding/unfolding of small α-helices, azobenzene derivatives have also been incorporated in small cyclic peptides[10,11] and β-hairpins.[12−15] More complex structural processes induced by that concept include the folding/unfolding transition of relatively large proteins,[16,17] the

Received: March 25, 2019
Published: June 11, 2019

DOI: 10.1021/jacs.9b03222
J. Am. Chem. Soc. 2019, 141, 10702−10710

modulation of enzymatic activity,[18−20] cell−cell adhesion,[21] or ion-channel activity.[3] *In vivo* applications became possible with the design of biocompatible azobenzene derivatives, e.g., in zebrafish embryos[22,23] or to control mitosis.[24]

In this study, we modulate the binding affinity within the noncovalent ribonuclease S complex (RNase S) with the ultimate goal of achieving photoinduced peptide unbinding. RNase S results from limited, site-specific hydrolysis of ribonuclease A, an enzyme from bovine pancreas.[25] Under controlled conditions, subtilisin can cleave a single peptide bond in RNase A and yields the RNase S complex composed of the S-peptide (residues 1−20) and the S-protein (residues 21−124).[25] Full enzymatic activity is restored when the two components form a native-like complex. The three-dimensional structure of the complex is essentially the same as that of intact RNase A, except from small differences near the cleaved peptide bond.[26] Consequently, the S-peptide adopts an α-helical structure when it is bound to the S-protein, while it is essentially disordered when isolated in solution due to its short sequence.[27] This property makes RNase S an excellent model to study the question whether the recognition mechanism between the S-protein and S-peptide can be characterized as "induced-fit" or as "conformational selection".[28−31] Furthermore, it opens the possibility to phototrigger binding/unbinding of the S-peptide, adopting the concept that has previously been used to control the α-helical content of isolated peptides via cross-linking two sites of the helix with an azobenzene moiety.[5−8] Figure 1 shows the construct we designed for this work: a S-peptide (blue) is bound to a S-protein (yellow), the former of which being photoswitchable and designed in a way that the azobenzene-moiety (red) in its *trans*-configuration destabilizes the α-helical content of the S-peptide.

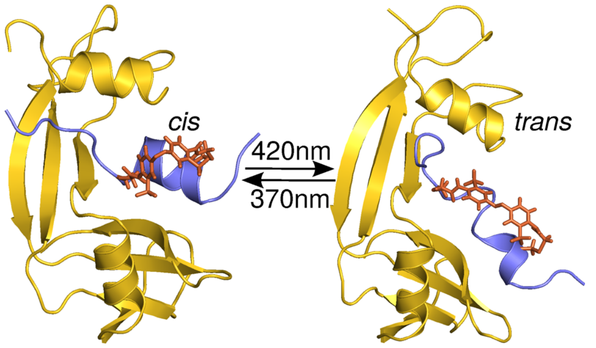


Figure 1. Molecular construct studied in this work: S-protein (yellow) with S-pep(6,13) (blue) bound to it in the two states of the photoswitch (orange).

There have been attempts to photocontrol RNase S in the past, mainly with the focus on influencing the enzymatic activity. Pioneering studies have been performed by Lui et al.[32] as well as Hamachi et al.[33] They created different mutants with the non-natural amino acid phenylazophenylalanine placed at different positions of the S-peptide. In contrast to our construct, the photoswitch does not cross-link two sites. Nonetheless, by photoswitching the configuration of the artificial side chain, particularly at position 13 of the S-peptide,[33] the enzymatic activity in one of the photoswitch states was reported to cease completely. In the subsequent study of Woolley's group,[34] the same model system has been studied more extensively. They could not reproduce the results from ref 33; rather, they found only a 2-fold difference in the activity between *cis* and *trans* states. In this study, binding affinities were also estimated, albeit indirectly via the enzymatic activity results. The largest modulation they found for the binding affinity in the *cis* vs the *trans* state was a factor 5. A different protein system that, however, uses a very similar design idea as our present one has been reported by Kneissl et al.,[35] who designed photocontrollable α-helical model peptides that bind to the Bcl-xl protein and are relevant for apoptosis. Up to 20-fold difference in binding affinity has been reported upon *trans* to *cis* isomerization of the photoswitch, but specific binding was preserved in both states with binding affinities in the nanomolar range.

2. DESIGN CRITERIA

For the design of the photoswitchable S-peptides, the following basic criteria were considered:

- A water-soluble derivative of azobenzene has been used as photoswitch, BSBCA (3,3′-bis(sulfonato)-4,4′-bis-(chloroacetamido) azobenzene).[36]
- A pair of amino acid residues needs to be replaced by cysteines for the site-selective linking of BSBCA. When choosing these sites, their effect for binding to the S-protein and/or the stability of the α-helix needs to be considered.
- Steric hindrance between cross-linked photoswitch and S-protein binding cleft should be avoided so that specific binding is preserved. To that end, the PyMOL software package[37] was used for visual examination and discrimination between solvent-exposed and nonsolvent-exposed amino acid residues of the S-peptide.
- The goal has been to stabilize the α-helical conformation in the more compact *cis* state of BSBCA, and to destabilize it in the *trans* state. This criterion determines the distance between sites that are being cross-linked.[9] To that end, we assumed an ideal length of BSBCA in its *cis* state of 11−16 Å,[17] and determined distances from the X-ray structure of RNase A (PDB ID: 2E3W).[38]

Within that framework, five different photoswitchable S-peptide variants have been designed. Ala6 has been chosen as the first anchoring point in all but one S-peptide variants, since it is the first amino acid residue (starting from N-terminus) that is part of the α-helical region and at the same time is solvent-exposed in the complex. Furthermore, it has been shown to have no effect on the binding affinity or the structure of the RNase S complex.[28,31,39]

To obtain a prototype S-peptide that preserves binding, we started with Ser15 as a second anchoring point, which also is known to be a residue that does not affect the binding.[28] Furthermore, it fulfills all aforementioned steric criteria; in particular, the end-to-end distance of the photoswitch in the *cis*-state fits the distance in the X-ray structure the best. We will denote that variant as S-pep(6,15). However, we learned quickly that this particular variant does not modulate the binding affinity significantly upon photoswitching, and we moved toward the next S-peptide variant, investigating the effect of a mutation that destabilizes the S-protein/S-peptide complex. We selected His12 for this purpose, whose mutation to Phe12 decreases the binding affinity by about a factor 25;[29] we will call that variant S-pep(6,15)H12F.

Furthermore, assuming that the α-helical content determines the binding, we considered S-pep(6,13) and S-pep(6,10) with shorter spacings 7 and 4, respectively, between the anchoring groups. According to ref 9, these spacings maximize the difference in stability of an α-helix with the photoswitch in the *cis* vs the *trans* states. Finally, we explored one variant in which we varied the starting point by one amino acid, S-pep(7,11), again with spacing 4. The amino acids that are affected in this case, Lys7 and Gln11, were both shown to play a role in the stabilization of the helix and in complex formation.[40,41]

3. MATERIALS AND METHODS

3.1. Preparation and Purification of S-Protein and S-Peptide Variants. The S-protein was prepared from the commercial bovine ribonuclease A (Sigma-Aldrich) by limited proteolysis with subtilisin, in essence as described in refs 25, 42. The proteolysis mixture was then applied to size-exclusion chromatography on a Sephadex G-75 in order to isolate the S-protein from the S-peptide as well as from subtilisin and residual unhydrolized RNase A. Wild-type S-peptide (sequence KETAAAKFERQHMDSSTSAA) and four modified S-peptide mutants with pairs of cysteine residues on different positions (i.e., S-pep(6,15), S-pep(6,13), S-pep(6,10) and S-pep(7,11), see Section 2) were synthesized by standard fluorenylmethoxycarbonyl(Fmoc)-based solid-phase peptide-synthesis on a Liberty 1 peptide synthesizer (CEM Corporation, Matthews, NC, USA). For S-pep(6,15)H12F, an additional His-to-Phe mutation was introduced at position 12. All amino acids were purchased from Novabiochem (La Jolla, CA, USA). The peptides were subsequently purified using reverse-phase C18 high performance liquid chromatography (HPLC) with an acetonitrile gradient 0−100% in 10 column volumes. The cysteine residues were subsequently cross-linked with the photoisomerizable linker BSBCA,[36] and purified again with the same conditions as in the previous step. The purity of S-protein and all S-peptide variants was confirmed by mass-spectrometry. Solutions of S-protein and S-peptide variants were prepared in 50 mM sodium phosphate buffer at pH 7.0. Concentrations were determined by amino acid analysis.

The *trans*-configuration is the thermodynamically more stable one and accumulates to essentially 100% upon thermal relaxation, which happens on a time scale of 1−10 h (see Supporting Information, Figure S1A). Typically, we left the sample in the dark overnight before an experiment of the *trans*-configuration was performed. To prepare the *cis*-configuration, we illuminated the sample with a 370 nm cw diode laser (90 mW, CrystaLaser). Its power is sufficient to switch a sample in minutes, and we verified that >85% *cis* is prepared in that way (see Figure S1B and its discussion).

3.2. Isothermal Titration Calorimetry (ITC). The ITC measurements were performed on a MicroCal iTC200 from Malvern (Malvern, UK) at 25 or 21 °C for wild-type and photoswitchable S-peptides, respectively. The sample cell contained 250 μL of S-protein solution at concentrations that varied according to the expected affinity. 10× higher S-peptide concentration was titrated into the sample cell. The first injection was 0.4 μL, and subsequent injections every 120 s were 2.0 μL. We verified that ITC reveals reliable results for the binding affinity also in the *cis*-state, despite the fact that the overall measurement time is comparable to that of the thermal *cis*-to-*trans* back-isomerization for the free peptide, and despite the fact that this reaction releases heat[43] that is comparable to the heat of binding (see Figure S1A and its discussion). K_d values were determined from the ITC data using the instrument's software and assuming one binding site.

3.3. Circular Dichroism (CD) Measurements. The CD measurements were performed on a Jasco (Easton, MD) model J-810 spectropolarimeter in a 0.1 cm quartz cuvette. We made sure the measurement was fast enough to prevent a significant amount of thermal back-isomerization, and that the intensity of the light used by the CD spectrometer was low enough so the amount of molecules that photoisomerized due to that light was negligible (see Figure S2 in Supporting Information).

3.4. Intrinsic Tyrosine Fluorescence Quenching. Fluorescence measurements were performed using two different fluorimeters. A commercial instrument (PerkinElmer) was used for the wild-type S-peptide, setting excitation and emission wavelengths to 285 and 306 nm, respectively. A larger sensitivity was needed for the photoswitchable variants due to the large absorption of the photoswitch itself, which furthermore changes between *cis* and *trans* states, and hence required an extremely low concentration to avoid fluorescence reabsorption (optical densities were kept below OD = 0.05). At the same time, the excitation intensity needed to be very low to avoid inducing isomerization of the photoswitch during the course of the measurement (see Supporting Information for details). The larger sensitivity was achieved with a home-built fluorimeter, which employed 266 nm for the excitation obtained from the third harmonic of a Ti:sapphire amplified laser system (Spectra Physics, Spitfire). The fluorescence was detected in a 90° geometry using a single photon counter (PMA 175-N-M, PicoQuant) together with a special filter (XRR0340 Asahi Spectra USA) to remove both the 266 and 370 nm radiations needed to promote the fluorescence of the tyrosines and the *trans* to *cis* isomerization of the photoswitch, respectively.

3.5. Molecular Dynamics (MD) Simulations. To generate initial structures of RNase S for the MD simulations, we started with the RNase complex structure (PDB ID: 1Z3P),[44] mutated Nva13 to Met13, and then attached five additional residues from an X-ray structure of RNase A (PDB ID: 2E3W)[38] at the C-terminus of the S-peptide. Following the procedure from ref 45, we next attached the azobenzene photoswitch on cysteine side chains by mutating the respective residues *i*, *j* of S-peptides according to experimental labeling positions (see Section 2). All MD simulations were performed using the GROMACS software package v4.6.7 (ref 46) with hybrid GPU-CPU acceleration scheme, the Amber03ws force field[47] and TIP4P-2005 water[48] solvent molecules. After energy minimization, all systems were equilibrated for 10 ns, and simulated at 300 K for at least 400 ns (see Supporting Information for details). Moreover, a MD simulation of wild-type RNase S was performed, in order to show the similarity of the conformational distributions of cross-linked vs wild-type peptides (Figure S3).

MD trajectories were preprocessed by performing a principal component analysis on backbone dihedral angles (dPCA+)[49] of residues 3 to 13 of S-peptides. Including only the first few principal components with 80% of the total sum of eigenvalues (depending on the individual system, about five to six), we define a reduced coordinate space of significantly lower dimensionality.[50] After dimensionality reduction, robust density-based clustering[51] was performed to identify the metastable conformational states of the individual systems (see Supporting Information for details). On the basis of most populated states, representative structures and respective flexibilities were obtained by employing *gmx rmsf* in Gromacs and displayed via *PyMOL*.[37]

4. RESULTS

4.1. Experimental Results. We used three independent methods to determine the binding affinity of the various S-peptides to the S-protein: ITC, CD spectroscopy and intrinsic tyrosine fluorescence quenching. Each one of these methods focuses on different aspects of binding, and has its strengths and weaknesses. For some of the S-peptides, more than one method could be applied with reasonable confidence, allowing us to cross-validate the results.

Starting with the "gold standard", ITC,[52] it was possible to determine binding affinities for relatively strong binders, see Figure 2 and Table 1. Due to the limited solubility of both the S-protein and the S-peptide, we could not determine the binding affinity of S-pep(6,13) in either state of the photoswitch, and, more severely, also not for the *trans*-configuration of most of the other S-peptides, with the one

DOI: 10.1021/jacs.9b03222
J. Am. Chem. Soc. 2019, 141, 10702−10710

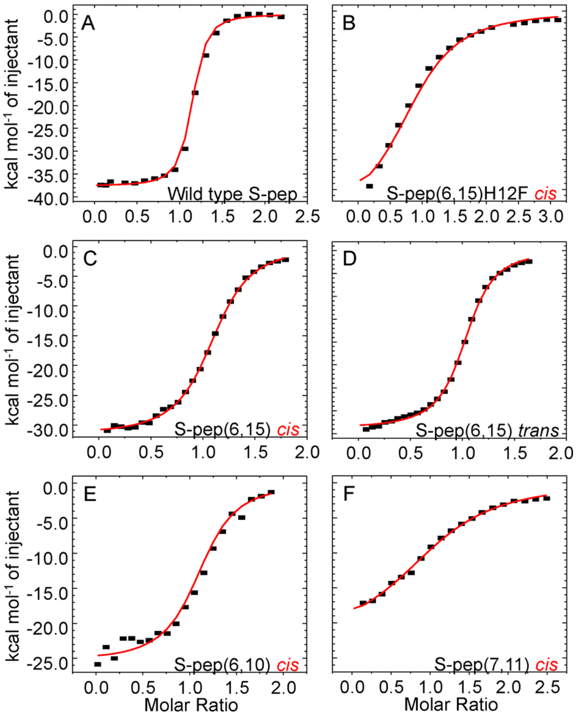


Figure 2. ITC thermograms of the S-protein and different S-peptide variants. (A) Wild-type S-peptide, (B) S-pep(6,15)H12F-*cis*, (C) S-pep(6,15)-*cis*, (D) S-pep(6,15)-*trans*, (E) S-pep(6,10)-*cis* and (F) S-pep(7,11)-*cis*.

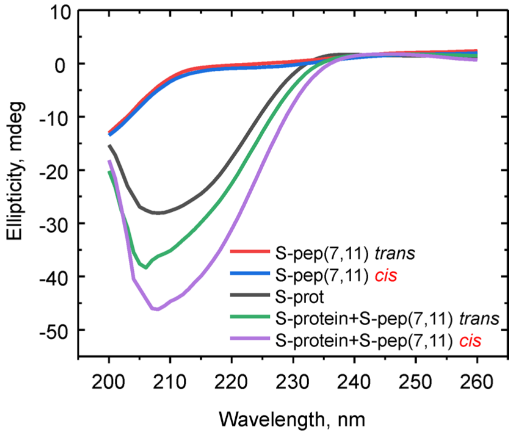


Figure 3. CD spectra of the S-peptide alone in *trans* (red) and *cis* (blue), the S-protein alone (black), and the S-protein with S-peptide in the *trans* (green) and *cis* state (purple), exemplified here for S-pep(7,11).

exception of S-pep(6,15), for which the effect of switching is negligible.

We therefore turn to CD spectroscopy in the next step, as our starting assumption has been that the photoswitch in the *trans*-configuration perturbs the helical conformation of the S-peptide, and CD spectroscopy is sensitive to exactly that structural aspect. To set the stage, we show in Figure 3 CD spectra of the S-pep(7,11) alone in its two states (red and blue), of the S-protein alone (black), as well as of mixtures of both at a concentration where the amount of binding has been ≈25% in the *trans*-state and ≈75% in the *cis*-state (*vide infra*). For the S-peptide alone, the CD spectra indicate a predominantly random-coil conformation, regardless of the state of the photoswitch (see Figure 3, red and blue). This finding is in agreement with the well-known fact that the S-peptide is disordered when it is free in solution.[29] The structural constrain of the photoswitch does not lead to any significant stabilization of the helical structure in the isolated form. Once bound to S-protein, however, the S-peptide adopts a helical structure, as deduced from a CD signal (Figure 3, green and purple) that is significantly larger than the sum of the signals from the S-peptide alone (red and blue) and S-protein alone (black). Furthermore, the results show that there are significant differences in helical content between the *cis* and *trans* states, indicating a larger amount of binding due to a higher affinity in the *cis*-state.

Besides folding of the S-peptide upon binding, the contribution of the S-protein *per se* to the overall CD signal might change as well. That can be seen from the fact that the difference between, e.g., the purple line in Figure 3 (S-protein plus ≈75% of bound S-pep(7,11)-*cis*) and the black line (S-protein alone), is much bigger than what the size of the fragments would suggest (20 amino acids for the S-peptide vs 104 amino acids of the S-protein). Indeed, it has been suggested that the S-protein is much more flexible in the absence of the S-peptide,[53] which is also the reason for difficulties in obtaining a crystal structure. Along these lines, a recent MD simulation has shown that helix II of the S-protein unfolds when the S-peptide unbinds.[31]

To determine the binding affinity, the CD signal at 225 nm[54] was measured as a function of S-peptide concentration, keeping the concentration of the S-protein constant at around 50 μM, see Figure 4. In these plots, "0%-fraction-bound" corresponds to the CD signal calculated as a trivial sum of the

Table 1. Dissociation Constants K_d (in μM) of the Various S-Peptide Variants and the S-Protein As Determined by ITC, CD and Intrinsic Tyrosine Fluorescence Quenching

	ITC		CD		fluorescence	
	cis	*trans*	*cis*	*trans*	*cis*	*trans*
S-pep(6,15)	1.5 ± 0.2	1.2 ± 0.8	1.7 ± 1.0	2.9 ± 0.8	3.6 ± 0.9	3.2 ± 1.3
S-pep(6,15)H12F	1.9 ± 0.4	–	1.5 ± 0.5	21 ± 1.8	–	–
S-pep(6,13)	–	–	70 ± 20	*[a]	130 ± 30	*[b]
S-pep(6,10)	0.6 ± 0.2	–	2.3 ± 1.2	47 ± 10	–	–
S-pep(7,11)	5.9 ± 0.2	–	6.2 ± 2.9	125 ± 15	–	–
wild-type	0.11 ± 0.01		0.14 ± 0.04		0.14 ± 0.02	

[a] No binding was observed. [b] Unspecific binding was observed, however, since the saturation plateau could not be reached, the binding constant could not be determined.

DOI: 10.1021/jacs.9b03222
J. Am. Chem. Soc. 2019, 141, 10702−10710

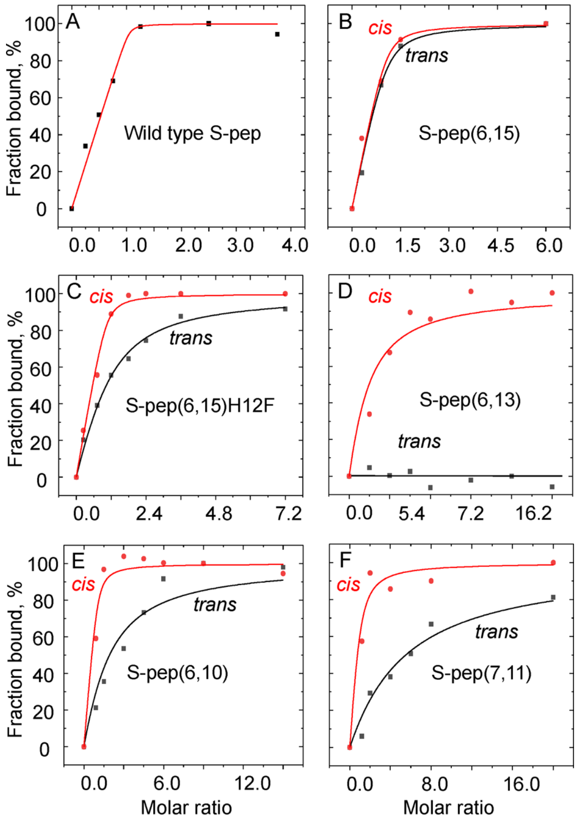


Figure 4. CD binding curves for various S-peptide variants. (A) Wild-type, (B) S-pep(6,15), (C) S-pep(6,15)H12F, (D) S-pep(6,13), (E) S-pep(6,10) and (F) S-pep(7,11). Measurements with the photoswitch in *cis* configuration are plotted in red, those in the *trans* configuration in black.

contributions from S-peptide alone and S-protein alone, and "100%-fraction-bound" to the plateau value reached at high enough S-peptide concentrations. To determine the binding affinity K_d, the data were fit to a chemical equilibrium, $P + L \rightleftharpoons PL$. Table 1 summarizes the results. It turns out that the method is more reliable for weak binders and as such complementary to ITC. That is, if the binding affinity becomes too large with K_d values much smaller than the concentration of the S-protein, the binding curve adopts a kink-like behavior (see in particular the wild-type data in Figure 4A) and the fit parameters become very insensitive to K_d; in essence, it is then a single data point at the kink that determines K_d. Fortunately, both ITC and CD work reasonably well for S-pep(6,15), allowing us to cross-validate the two methods.

Finally, fluorescence quenching was employed as a method to determine the dissociation constants for the *cis* and the *trans* forms of some of the peptides. The intrinsic tyrosine S-protein fluorescence decreases by ca. 20% upon binding in all cases. Local structural aspects like changes of the chemical environment via hydrogen bonds, disulfide bridges formation/breaking and resonance energy transfer mechanism can affect the intrinsic tyrosine fluorescence quenching of RNase S.[55,56] The obtained binding curves are shown in Figure 5, where the values for "0%-fraction-bound" and "100%-fraction-bound" have been determined in the same way as for the CD data. Wild-type and S-pep(6,15) were used to cross-validate the fluorescence method with CD and ITC techniques, with both the commercial and the home-built fluorimeters, respectively, and S-pep(6,13) was then measured as an example with very low binding affinity. The extracted binding affinities are summarized in Table 1. Overall, the results from three different methods are in good agreement with each other within experimental error.

4.2. Computational Results.

To study the relation between changes in binding affinity (i.e., experimental K_d obtained by CD spectroscopy) and secondary structural changes of the S-peptides, all-atom MD simulations were performed as described in Methods. As an overview, Figure 6 shows the resulting dynamics of secondary structures of the various S-peptides obtained from DSSP analysis.[57,58] In particular, we focus on the differences of helicity when changing from *cis* to *trans* configuration. In the case of S-pep(6,15) and S-pep(6,15)H12F, we find that the core region formed by residues 4−11 exhibits a stable α-helical conformation in both *cis*- and *trans*-configuration. Recalling that an α-helix of residues Thr3-Met13 is the hallmark of the native, active RNase S,[29] we conclude that the structures of these S-peptides are hardly perturbed by the photoswitch. This conclusion is also supported by the fact that the wild-type peptide exhibits about the same helicity as S-pep(6,15) and S-pep(6,15)H12F (see Figure S3). S-pep(6,13), on the other hand, is significantly more affected by *cis*-to-*trans* photoswitching. Apart from infrequent short perturbations, S-pep(6,13)-*cis* reveals a relatively stable α-helical structure, while its *trans* form mostly exist as a 3_{10} helix or random coil. In a similar way, S-pep(6,10) and S-pep(7,11) exhibit decreased α-helicity upon *cis*-to-*trans* photoswitching, although

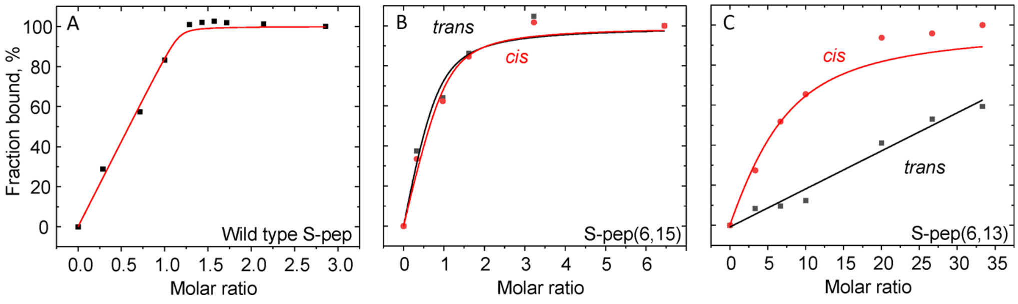


Figure 5. Fluorescence binding curves for various S-peptide variants. (A) Wild-type, (B) S-pep(6,15) and (C) S-pep(6,13). Measurements with the photoswitch in *cis* configuration are plotted in red, those in the *trans* configuration in black. In the case of S-pep(6,13)-*trans*, the saturation plateau was not reached; it was assumed to be the same as for S-pep(6,13)-*cis* for the purpose of plotting (but that does not have to be correct since unspecific binding implies that more than one S-peptide might bind). The linear fit in this case is displayed to guide the eyes.

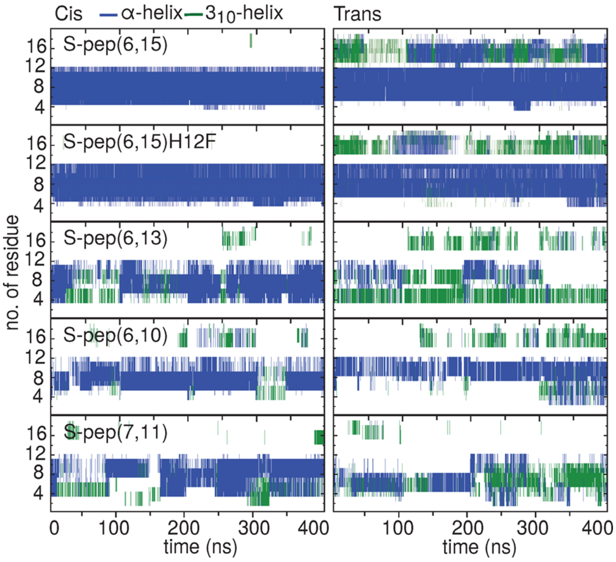


Figure 6. Time evolution of the S-peptides' secondary structure in *cis* (left) and *trans* (right) configuration, obtained from all-atom MD simulations after suitable equilibration (see Methods). α-Helix content is shown in blue, 3_{10} helix in green. Unstructured parts are shown in white.

the structural changes found in *trans* are less prominent compared to S-pep(6,13). Performing a time average over the trajectories, the above findings can be represented by a probability distribution of the α-helicity of the various systems (see Figure S3, Supporting Information).

To connect these computational results for the α-helicity α_H to the experimental findings of K_d in form of a quantitative structure activity relationship (QSAR), we used a Boltzmann statistics-based ansatz, which assumes that the free energy of binding, $\Delta G \sim \ln(K_d)$, is linearly dependent on the number of peptide residues in α-helical conformation. This gives

$$\alpha_H = -c_\alpha k_B T \ln(K_d) + \alpha_{H_0} \qquad (1)$$

where c_α accounts for the free energy gain per helical residue $\Delta G = \alpha_H / c_\alpha$, and α_{H_0} is a reference peptide helicity. As shown in Figure 7, this ansatz reveals a convincing fit ($R^2 = 0.85$) for the *trans*-state. The fit is less clear ($R^2 = 0.65$) for the *cis*-state, where we also added the results of the wild-type. The resulting fit parameters for *trans* ($c_\alpha = 3.4 \pm 0.9$, $\alpha_{H0} = 53 \pm 6$) and *cis* ($c_\alpha = 4.2 \pm 2.5$, $\alpha_{H0} = 44 \pm 6$) agree with each other within their error ranges.

Figure 8 combines these results and correlates the calculated change in α-helicity $\Delta\alpha_H = \alpha_{H(cis)} - \alpha_{H(trans)}$ to the experimentally measured change in affinity $\Delta K_d = K_{d(trans)} - K_{d(cis)}$. In nice agreement, both simulated $\Delta\alpha_H$ and experimental ΔK_d show an ascending behavior starting with S-pep(6,15) via S-pep(6,15)H12F and S-pep(6,10) to S-pep(7,11) and S-pep(6,13).

It is interesting to connect the above-discussed helicity changes with the prevailing molecular structures of the S-peptides. To this end, robust density-based clustering[51] (see Methods) was performed for all systems and the resulting main conformational states were analyzed in detail (see Figures S4 and S5, Supporting Information). Overall, we find that all S-peptides can be well described by only four conformational states, where the dominant state typically occurs with 40−70% population. Figure 9 shows the resulting main structure of all S-peptides (the S-protein is not shown for clarity). In line with the discussion of Figure 6, we find that *cis*-to-*trans* photoswitching has the least impact on the secondary structures of S-pep(6,15) and S-pep(6,15)H12F, while the most significant changes are found for S-pep(6,13), which essentially loses its α-helical conformation. Coloring residues according to their fluctuations, we find that the structures are well-defined in both *cis* and *trans* and exhibit increased flexibility only at the terminals. That is, even the apparently disordered regions of S-pep(6,13) and S-pep(7,11) reveal stable structures.

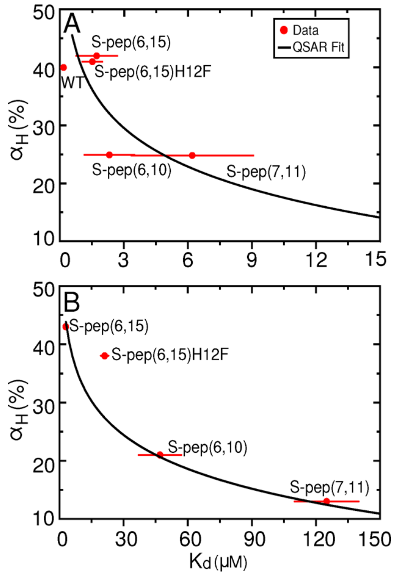


Figure 7. (A) α-Helicity vs experimental K_d in *cis*-configuration, together with the wild-type (WT), and fit obtained by eq 1. (B) The same for the *trans*-configuration.

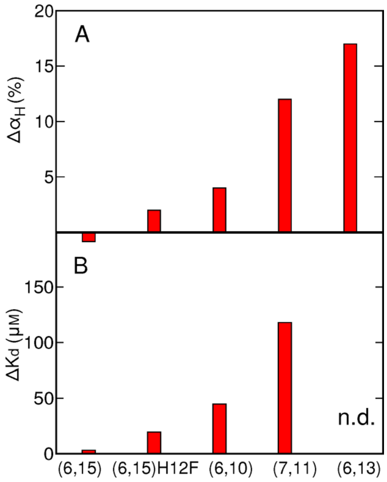


Figure 8. (A) Change in α-helicity difference: $\Delta\alpha_H = \alpha_{H(cis)} - \alpha_{H(trans)}$ and (B) change in experimental $\Delta K_d = K_{d(trans)} - K_{d(cis)}$, as determined from CD spectroscopy (Table 1). No experimental data exist for K_d in the *trans*-state of S-pep(6,13), which is why no bar is shown in panel B, but it is clear from Figure 4D that the difference is large.

DOI: 10.1021/jacs.9b03222
J. Am. Chem. Soc. 2019, 141, 10702−10710

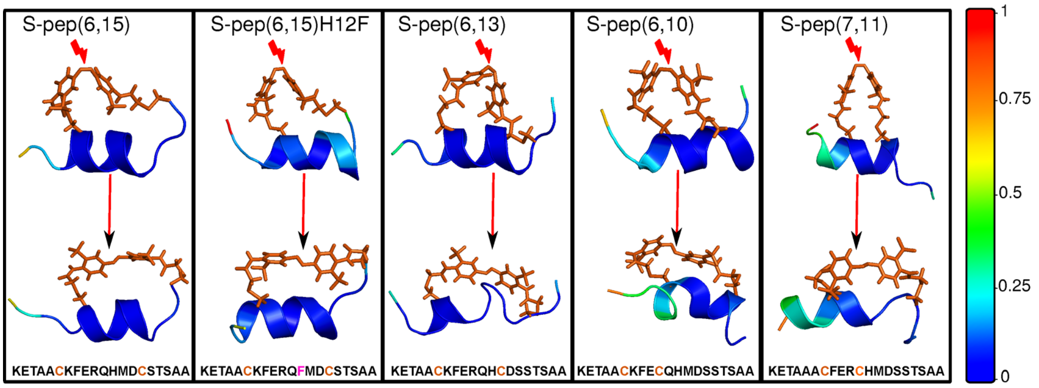


Figure 9. Most populated structures obtained from a clustering analysis, see text and Supporting Information for details. The S-protein is not shown for clarity. The residues are colored according to normalized root-mean-square fluctuations.

5. DISCUSSION AND CONCLUSION

To facilitate photocontrolling of peptide–protein binding, we have designed various photoswitchable S-peptides, switching their helical content, and measured their binding affinities to the S-protein using a variety of experimental techniques (Table 1). α-Helices are structural motifs that very commonly are relevant for protein–protein, protein–peptide and protein–DNA interactions and therefore represent important targets for modulation of binding affinities. Here, we have introduced a concept by which the helical content can be modulated in a very controlled manner. The overall disruption of secondary structures depends not only upon the distance between anchor residues of photoswitch but also the exact position of the photoswitch.

By performing MD simulations of the *cis* and *trans* configurations of the RNase S complex, we were able to relate simulated α-helicities α_H to experimentally obtained binding affinities K_d, which are in surprisingly good agreement (Figure 8). Using the simple QSAR ansatz of eq 1, we have shown that the α-helicity is inversely related to the experimental K_d values (Figures 7). This suggests that relatively short (sub-μs) MD simulations are indicative of the long-time binding or dissociation behavior of the RNase S complex. The secondary structure analysis of the S-peptides in Figure 9) shows a significant decrease in α-helicity upon *cis*-to-*trans* isomerization for all systems, with the one exception of S-pep(6,15). The latter is in line with the findings of ref 9 studying the photoswitching of isolated helical peptides, which revealed that the (i,i+9) spacing in S-pep(6,15) is the dividing point between decreasing or increasing the α-helical content upon *cis*-to-*trans* isomerization.

The largest change in α-helicity is found for S-pep(6,13) and S-pep(7,11) with spacings (i,i+7) and (i,i+4), respectively. The corresponding structures of the *trans* configurations in Figure 9 reveal a considerable loss of α-helical conformation. In particular, S-pep(6,13)-*trans* adopts a helical conformation only at its N-terminus, while its C-terminus is completely disordered. It has been suggested that the N-terminal part of the S-peptide is unzipping from the S-protein binding groove due to equilibrium state fluctuations.[31] Furthermore, based on an alanine mutant screening approach,[31] it has been shown that the N-terminus (more precisely residues 1–7) contributes to binding only to a minor extent. Hence, even though a small percentage of helicity was preserved in the case of S-pep(6,13)-*trans*, no specific binding is observed, since the residual helicity is located in relatively unimportant part for the binding.

It is interesting to note that the isomerization of the photoswitch does not lead to a significant increase of peptide fluctuations. Instead, Figures 9 and S4 (see Supporting Information) suggest that peptide structures are mostly well-defined in both the *cis* and the *trans*-state and exhibit fluctuations only at the termini. Especially the parts of the peptide that lie between the clamping points of the photoswitch display highly stable structures, even in apparently disordered states. Furthermore, in the case of S-pep(6,13), the peptide structure is more stable in the disordered *trans* state than in helical *cis* state, in clear contrast to intuition. This stabilization of disordered states turns out to be one of the most interesting outcomes of our photolabeling strategy. In combination with structural predictions from MD simulations, it allows a rather distinct manipulation of the peptide structure depending on the photoswitch attachment points.

We consider S-pep(6,13) the most interesting result in our series of mutants. First, while no binding could be detected for S-pep(6,13)-*trans* via CD spectroscopy (Figure 4D), some degree of binding is observed by fluorescence quenching (Figure 5C), illustrating the different aspects of binding to which the two methods are sensitive. That is, while CD measures helical content, fluorescence quenching measures spatial proximity. The difference between CD and fluorescence data can probably be explained by the fact that binding is unspecific in this case, i.e., more in the sense of aggregation, and hence does not induce any α-helical structure. Since we did not reach a plateau in fluorescence intensity in this case, we cannot truly quantify the binding affinity, but it is probably in the mM range. And indeed, such weak binding can typically no longer be characterized as "specific", in particular if the binding partner is a floppy peptide like in the present case.[59]

Second, the largest effects observed in the literature for photoswitchable protein–peptide complexes are K_d ratios of about 20 (ref 35), and the effect is equally large for S-pep(6,10) and S-pep(7,11). For S-pep(6,13), the ratio is even larger, i.e., effectively infinite, as no specific binding is detectable by CD spectroscopy in the *trans*-state. To the best of our knowledge, such a big effect has not been achieved so far for any photoswitchable protein–peptide complex. This property makes the system particularly interesting for time-resolved studies, in which the light-driven isomerization of the photoswitch triggers unbinding of the S-peptide, and the

DOI: 10.1021/jacs.9b03222
J. Am. Chem. Soc. 2019, 141, 10702–10710

structural response of the protein is investigated by, e.g., transient IR spectroscopy or transient X-ray scattering experiments in upcoming free electron lasers.[1] Ideally, in such an experiment, the S-peptide should unbind as quickly as possible; only then the temporal ordering of unbinding event vs structural response could be measured, addressing the long-standing question whether the process can be described as "conformational selection" or as "induced fit". However, the unbinding rate constant k_{off} scales essentially as K_{d} (since k_{on} is limited by diffusion and as such essentially a constant), hence to speed up k_{off}, K_{d} in one state of the photoswitch should be as large as possible. Transient IR experiments are currently underway.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.9b03222.

> Experimental validation of ITC as a method to determine binding affinities also in the *cis*-state, amount of *cis* prepared by illumination at 370 nm, effect of the measurement light in CD and fluorescence spectroscopy, more detailed MD methods, description of the helicity and contacts of wild-type and labeled systems, as well as detailed results of the clustering analysis (Figures S4 and S5) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*peter.hamm@chem.uzh.ch
*stock@physik.uni-freiburg.de

ORCID

Steffen Wolf: 0000-0003-1752-6175
Gerhard Stock: 0000-0002-3302-3044
Peter Hamm: 0000-0003-1106-6032

Author Contributions

§Both authors contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Ilian Jelezarov for important discussions concerning the ITC, Rolf Pfister for the synthesis of the peptides and BSBCA, Ben Schuler and his group for their continuous help with the protein chemistry, and the Functional Genomics Center Zurich, especially Serge Chesnov for help with the mass spectrometry and Birgit Roth for amino acid analysis. The work has been supported in part by the Swiss National Science Foundation (SNF) through Grant 200021_165789, as well as by the Deutsche Forschungsgemeinschaft through Grant STO 247/10-2. We acknowledge support by the High Performance and Cloud Computing Group at the Zentrum für Datenverarbeitung of the University of Tübingen, the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant no INST 37/935-1 FUGG (RV bw16I016) and the Black Forest Grid Initiative.

■ REFERENCES

(1) Standfuss, J. Membrane protein dynamics studied by X-ray lasers – or why only time will tell. *Curr. Opin. Struct. Biol.* **2019**, *57*, 63–71.

(2) Hamm, P.; Helbing, J.; Bredenbeck, J. Two-Dimensional Infrared Spectroscopy of Photoswitchable Peptides. *Annu. Rev. Phys. Chem.* **2008**, *59*, 291–317.

(3) Zatsepin, T. S.; Abrosimova, L. A.; Monakhova, M. V.; Le, T. H.; Pingoud, A.; Kubareva, E. A.; Oretskaya, T. S. Design of photo-controlled biomolecules based on azobenzene derivatives. *Russ. Chem. Rev.* **2013**, *82*, 942–963.

(4) Beharry, A. A.; Woolley, G. A. Azobenzene photoswitches for biomolecules. *Chem. Soc. Rev.* **2011**, *40*, 4422–4437.

(5) Kumita, J. R.; Smart, O. S.; Woolley, G. A. Photo-control of helix content in a short peptide. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 3803–3808.

(6) Bredenbeck, J.; Helbing, J.; Kumita, J. R.; Woolley, G. A.; Hamm, P. alpha-Helix formation in a photoswitchable peptide tracked from picoseconds to microseconds by time resolved IR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 2379–2384.

(7) Woolley, G. A. Photocontrolling peptide alpha helices. *Acc. Chem. Res.* **2005**, *38*, 486–493.

(8) Ihalainen, J. A.; Paoli, B.; Muff, S.; Backus, E. H. G.; Bredenbeck, J.; Woolley, G. A.; Caflisch, A.; Hamm, P. Alpha-Helix folding in the presence of structural constraints. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9588–9593.

(9) Flint, D. G.; Kumita, J. R.; Smart, O. S.; Woolley, G. A. Using an Azobenzene Cross-Linker to Either Increase or Decrease Peptide Helix Content upon Trans-to-Cis Photoisomerization. *Chem. Biol.* **2002**, *9*, 391–397.

(10) Bredenbeck, J.; Helbing, J.; Sieg, A.; Schrader, T.; Zinth, W.; Renner, C.; Behrendt, R.; Moroder, L.; Wachtveitl, J.; Hamm, P. Picosecond conformational transition and equilibration of a cyclic peptide. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 6452–6457.

(11) Satzger, H.; Root, C.; Renner, C.; Behrendt, R.; Moroder, L.; Wachtveitl, J.; Zinth, W. Picosecond dynamics in water-soluble azobenzene-peptides. *Chem. Phys. Lett.* **2004**, *396*, 191–197.

(12) Aemissegger, A.; Krautler, V.; van Gunsteren, W. F.; Hilvert, D. A photoinducible beta-hairpin. *J. Am. Chem. Soc.* **2005**, *127*, 2929–2936.

(13) Jurt, S.; Aemissegger, A.; Güntert, P.; Zerbe, O.; Hilvert, D. A Photoswitchable Miniprotein Based on the Sequence of Avian Pancreatic Polypeptide. *Angew. Chem., Int. Ed.* **2006**, *45*, 6297–6300.

(14) Dong, S. L.; Loweneck, M.; Schrader, T. E.; Schreier, W. J.; Zinth, W.; Moroder, L.; Renner, C. A photocontrolled beta-hairpin peptide. *Chem. - Eur. J.* **2006**, *12*, 1114–1120.

(15) Schrader, T. E.; Schreier, W. J.; Cordes, T.; Koller, F. O.; Babitzki, G.; Denschlag, R.; Renner, C.; Löweneck, M.; Dong, S.-L.; Moroder, L.; Tavan, P.; Zinth, W. Light-triggered beta-hairpin folding and unfolding. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 15729–15734.

(16) Lorenz, L.; Kusebauch, U.; Moroder, L.; Wachtveitl, J. Temperature- and Photocontrolled Unfolding/Folding of a Triple-Helical Azobenzene-Stapled Collagen Peptide Monitored by Infrared Spectroscopy. *ChemPhysChem* **2016**, *17*, 1314–1320.

(17) Zhang, F.; Zarrine-Afsar, A.; Al-Abdul-Wahid, M. S.; Prosser, R. S.; Davidson, A. R.; Woolley, G. A. Structure-based approach to the photocontrol of protein folding. *J. Am. Chem. Soc.* **2009**, *131*, 2283–2289.

(18) Schierling, B.; Noel, A.-J.; Wende, W.; Hien, L. T.; Volkov, E.; Kubareva, E.; Oretskaya, T.; Kokkinidis, M.; Rompp, A.; Spengler, B.; Pingoud, A. Controlling the enzymatic activity of a restriction enzyme by light. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 1361–1366.

(19) Hoersch, D. Engineering a light-controlled F 1 ATPase using structure-based protein design. *PeerJ* **2016**, *4*, No. e2286.

(20) Eisel, B.; Hartrampf, F. W.; Meier, T.; Trauner, D. Reversible optical control of F1Fo-ATP synthase using photoswitchable inhibitors. *FEBS Lett.* **2018**, *592*, 343–355.

(21) Ritterson, R. S.; Kuchenbecker, K. M.; Michalik, M.; Kortemme, T. Design of a photoswitchable cadherin. *J. Am. Chem. Soc.* **2013**, *135*, 12516–12519.

(22) Beharry, A. A.; Wong, L.; Tropepe, V.; Woolley, G. A. Fluorescence imaging of azobenzene photoswitching in vivo. *Angew. Chem., Int. Ed.* **2011**, *50*, 1325–1327.

(23) Wachtveitl, J.; Zumbusch, A. Azobenzene: An Optical Switch for in vivo Experiments. *ChemBioChem* **2011**, *12*, 1169−1170.

(24) Borowiak, M.; Nahaboo, W.; Reynders, M.; Nekolla, K.; Jalinot, P.; Hasserodt, J.; Rehberg, M.; Delattre, M.; Zahler, S.; Vollmar, A.; Trauner, D.; Thorn-Seshold, O. Photoswitchable Inhibitors of Microtubule Dynamics Optically Control Mitosis and Cell Death. *Cell* **2015**, *162*, 403−411.

(25) Richards, F. M.; Vithayathil, P. J. The Preparation of Subtilisin-modified Ribonuclease and the Separation of the Peptide and Protein Components. *J. Biol. Chem.* **1959**, *234*, 1459−1465.

(26) Wlodawer, A.; Sjolin, L. Hydrogen exchange in RNase A: Neutron diffraction study. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79*, 1418−1422.

(27) Schreier, A. A.; Baldwin, R. L. Mechanism of Dissociation of S-Peptide from Ribonuclease. *Biochemistry* **1977**, *16*, 4203−4209.

(28) Goldberg, J. M.; Baldwin, R. L. A specific transition state for S-peptide combining with folded S-protein and then refolding. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 2019−2024.

(29) Bachmann, A.; Wildemann, D.; Praetorius, F.; Fischer, G.; Kiefhaber, T. Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 3952−3957.

(30) Dogan, J.; Gianni, S.; Jemth, P. The binding mechanisms of intrinsically disordered proteins. *Phys. Chem. Chem. Phys.* **2014**, *16*, 6323−6331.

(31) Luitz, M. P.; Bomblies, R.; Zacharias, M. Comparative Molecular Dynamics Analysis of RNase-S Complex Formation. *Biophys. J.* **2017**, *113*, 1466−1474.

(32) Liu, D.; Karanicolas, J.; Yu, C.; Zhang, Z.; Woolley, G. A. Site-specific incorporation of photoisomerizable azobenzene groups into ribonuclease S. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 2677−2680.

(33) Hamachi, I.; Hiraoka, T.; Yamada, Y.; Shinkai, S. Photo-switching of the Enzymatic Activity of Semisynthetic Ribonuclease S Bearing Phenylazophenylalanine at a Specific Site. *Chem. Lett.* **1998**, *27*, 537−538.

(34) James, D.; Burns, D. C.; Woolley, G. Kinetic characterization of ribonuclease S mutants containing photoisomerizable phenylazophenylalanine residues. *Protein Eng., Des. Sel.* **2001**, *14*, 983−991.

(35) Kneissl, S.; Loveridge, E. J.; Williams, C.; Crump, M. P.; Allemann, R. K. Photocontrollable Peptide-Based Switches Target the Anti-Apoptotic Protein Bcl-x L. *ChemBioChem* **2008**, *9*, 3046−3054.

(36) Zhang, Z.; Burns, D. C.; Kumita, J. R.; Smart, O. S.; Woolley, G. A. A Water-Soluble Azobenzene Cross-Linker for Photocontrol of Peptide Conformation. *Bioconjugate Chem.* **2003**, *14*, 824−829.

(37) *The PyMOL Molecular Graphics System*, Version 1.8; Schrödinger, LLC, 2015.

(38) Boerema, D. J.; Tereshko, V. A.; Kent, S. B. H. Total synthesis by modern chemical ligation methods and high resolution (1.1 Å) x-ray structure of ribonuclease a. *Biopolymers* **2008**, *90*, 278−286.

(39) Stelea, S. D.; Keiderling, T. A. Pretransitional Structural Changes in the Thermal Denaturation of Ribonuclease S and S Protein. *Biophys. J.* **2002**, *83*, 2259−2269.

(40) Dice, J. F.; Chiang, H.-l.; Spencer, E. P.; Backer, J. M. Regulation of Catabolism of Microinjected Ribonuclease A. *J. Biol. Chem.* **1986**, *261*, 6853−6859.

(41) Dice, J. F. Microinjected ribonuclease A as a probe for lysosomal pathways of intracellular protein degradation. *J. Protein Chem.* **1988**, *7*, 115−127.

(42) Bagchi, S.; Boxer, S. G.; Fayer, M. D. Ribonuclease S dynamics measured using a nitrile label with 2D IR vibrational echo spectroscopy. *J. Phys. Chem. B* **2012**, *116*, 4034−4042.

(43) Dias, A. R.; Minas Da Piedade, M. E.; Martinho Simoes, J. A.; Simoni, J. A.; Teixeira, C.; Diogo, H. P.; Meng-Yan, Y.; Pilcher, G. Enthalpies of formation of cis-azobenzene and trans-azobenzene. *J. Chem. Thermodyn.* **1992**, *24*, 439−447.

(44) Das, M.; Rao, B. V.; Ghosh, S.; Varadarajan, R. Attempts to delineate the relative contributions of changes in hydrophobicity and packing to changes in stability of ribonuclease s mutants. *Biochemistry* **2005**, *44*, 5923−5930.

(45) Zanobini, C.; Bozovic, O.; Jankovic, B.; Koziol, K. L.; Johnson, P. J. M.; Hamm, P.; Gulzar, A.; Wolf, S.; Stock, G. Azidohomoalanine: A Minimally Invasive, Versatile, and Sensitive Infrared Label in Proteins To Study Ligand Binding. *J. Phys. Chem. B* **2018**, *122*, 10118−10112.

(46) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to super-computers. *SoftwareX* **2015**, *1−2*, 19−25.

(47) Best, R. B.; Zheng, W.; Mittal, J. Balanced protein-water interactions improve properties of disordered proteins and non-specific protein association. *J. Chem. Theory Comput.* **2014**, *10*, 5113−5124.

(48) Abascal, J. L. F.; Sanz, E.; García Fernández, R.; Vega, C. A potential model for the study of ices and amorphous water: TIP4P/Ice. *J. Chem. Phys.* **2005**, *122*, 234511.

(49) Sittel, F.; Filk, T.; Stock, G. Principal component analysis on a torus: Theory and application to protein dynamics. *J. Chem. Phys.* **2017**, *147*, 244101.

(50) Sittel, F.; Stock, G. Perspective: Identification of collective variables and metastable states of protein dynamics. *J. Chem. Phys.* **2018**, *149*, 150901.

(51) Sittel, F.; Stock, G. Robust density-based clustering to identify metastable conformational states of proteins. *J. Chem. Theory Comput.* **2016**, *12*, 2426−2435.

(52) Cooper, M. A. Label-free screening of bio-molecular interactions. *Anal. Bioanal. Chem.* **2003**, *377*, 834−842.

(53) Rosa, J. J.; Richards, F. M. Hydrogen Exchange from Identified Regions of the S-Protein Component of Ribonuclease as a Function of Temperature, pH, and the Binding of S-Peptide. *J. Mol. Biol.* **1981**, *145*, 835−851.

(54) Labhardt, A. M. Kinetic circular dichroism shows that the S-peptide a-helix of ribonuclease S unfolds fast and refolds slowly. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 7674−7678.

(55) Cowgill, R. W. Fluorescence and the structure of proteins. III. Effects of denaturation on fluorescence of insulin and ribonuclease. *Arch. Biochem. Biophys.* **1964**, *104*, 84−92.

(56) Noronha, M.; Lima, J. C.; Paci, E.; Santos, H.; Maçanita, A. L. Tracking local conformational changes of ribonuclease A using picosecond time-resolved fluorescence of the six tyrosine residues. *Biophys. J.* **2007**, *92*, 4401−4414.

(57) Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577−2637.

(58) Baakman, C.; Krieger, E.; Black, J.; Touw, W. G.; Vriend, G.; Joosten, R. P.; te Beek, T. A. A series of pdb-related databanks for everyday needs. *Nucleic Acids Res.* **2015**, *43*, D364−D368.

(59) Schreiber, G.; Keating, A. E. Protein binding specificity versus promiscuity. *Curr. Opin. Struct. Biol.* **2011**, *21*, 50−61.

Supporting Information: Photocontrolling Protein-Peptide Interactions: From Minimal Perturbation to Complete Unbinding

Brankica Jankovic,[1,3] Adnan Gulzar[2,3], Claudio Zanobini,[1]
Olga Bozovic,[1] Steffen Wolf[2], Gerhard Stock,[2,4] Peter Hamm[1,4]
[1]*Department of Chemistry, University of Zurich, Zurich, Switzerland*
[2]*Biomolecular Dynamics, Institute of Physics, Albert Ludwigs University, Freiburg, Germany*
[3]*both authors contributed equally*
[4]*corresponding authors:* peter.hamm@chem.uzh.ch and stock@physik.uni-freiburg.de

(Dated: June 20, 2019)

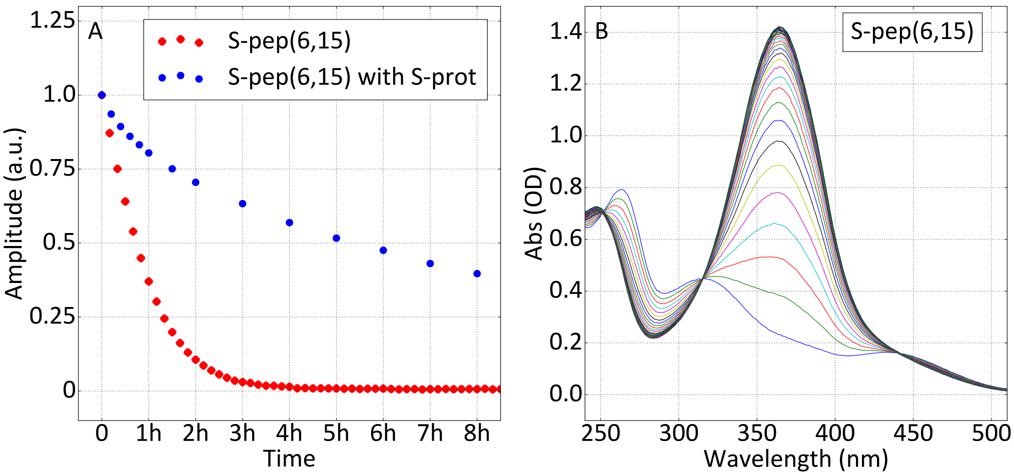


FIG. S1. (A) Thermal *cis*-to-*trans* back-relaxation exemplified for S-pep(6,15); in red for the isolated peptide in solution, and in blue for S-pep(6,15) bound to the S-protein. (B) Corresponding UV/VIS absorption spectra after 5 min illumination at 370 nm (blue line), preparing predominantly the *cis*-configuration, and subsequent thermal back-relaxation reestablishes the *trans*-band at 367 nm (spectra are recorded every 10 min). On the one hand, the kinetic data of (A) are retrieved from such a measurement series. On the other hand, the initial amount of *cis*-configuration (>85%) can be estimated from the ratio of absorbances at 370 nm (see text).

I. ITC OF S-PEPTIDES IN THE *CIS*-STATE

For the S-peptides in the *cis*-state, there has been a concern whether the binding affinity can be determined by ITC at all. In these experiments, the syringe was constantly illuminated with 370 nm, however, after injection of the photoswitchable peptide solution into the closed and dark sample cell, illumination was no longer possible and thermal back-isomerization of the photoswitch from *cis*-to-*trans* occurred to a certain extent. First, the time to complete the ITC measurement was kept as short as possible (45 min), which is in the same range as thermal back-relaxation of the free S-peptide (between 40-60 min depending on S-peptide variant, see Fig. S1A, red). We however found that thermal back relaxation is significantly longer once an S-peptide is bound to the S-protein (see Fig. S1A, blue), so the time that really counts is k_{on}, i.e. the time to bind right after an injection, which is short enough to be unproblematic.

Independent of that issue, thermal back-isomerization releases significant amount of heat, and one might expect a large background signal in the ITC data. In fact the heat released by back-isomerization[1] is about the same as the heat of binding. However, back-isomerization occurs continuously, in contrast to the binding-associated heat that occurs abruptly following every injection, and high-pass filtering in the ITC instrument can suppress that background completely. As a result, the binding affinities could be measured reliably also in the *cis*-state of the S-peptide, as our cross-validation with other methods has shown (see main text).

II. ESTIMATE OF THE AMOUNT OF *CIS*-CONFIGURATION

The *cis*-configuration is prepared by illuminating the sample with a 370 nm cw diode laser (CrystaLaser). Provided that the power of the light source is high enough to outperform thermal *cis*-*trans* back reaction (a condition which is safely fullfilled with that light source), this leads to a relatively high content of *cis*-configuration, since the *trans*-configuration has a strong band at this frequency position, while the absorption cross section of the *cis*-configuration is small. Fig. S1B (blue) shows an absorption spectrum of the *cis*-configuration prepared in that way, which then relaxes back to the *trans*-configuration. If we assume that the absorption cross section of the *cis*-configuration is exactly zero, we can estimate from the ratio of absorbances at 370 nm that the sample contains 85% *cis*-configuration. Since the absorption cross section of the *cis*-configuration is probably not quite zero, that estimate is considered a worst case scenario, i.e., a lower limit.

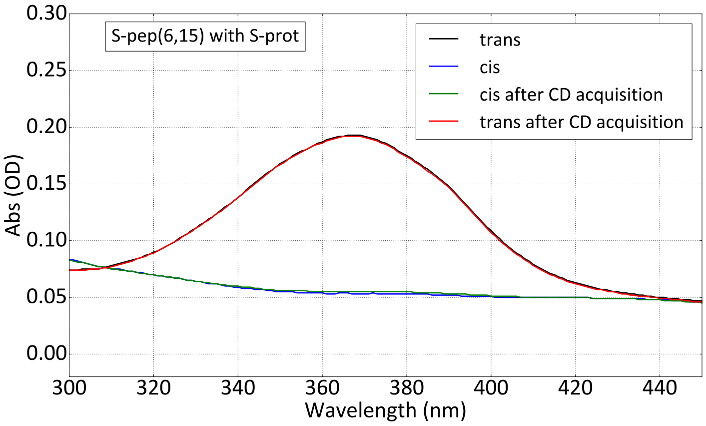


FIG. S2. UV absorption spectra of S-pep(6,15) in the *trans*-configuration before (black) and after (red) measuring the actual CD data that led to Fig. 4B, and the same for the *cis*-configuration in blue and green, respectively.

III. AMOUNT OF PHOTOISOMERIZATION DUE TO THE MEASUREMENT LIGHT FOR CD AND FLUORESCENCE SPECTROSCOPY

The light used to measure either the CD of the protein, or the intrinsic tyrosine fluorescence quenching, will also excite higher excited states of the azobenzene moiety, and induce its photoisomerization. We carefully checked that the accumulated intensity of that light during the measurement was low enough so that this effect is negligible. In case of CD spectroscopy, the corresponding spectrometer could also measure UV absorption spectra of the 370 nm band of the azobenzene moiety, which is very strong in the *trans*-configuration and essentially absent in *cis*. Fig. S2 shows spectra in that spectral range for the *trans*-configuration before (black) and after (red) measuring the actual CD data that led to Fig. 4B for S-pep(6,15), while the blue and green lines show the same for the *cis*-configuration. In neither case, any significant isomerisation could be detected.

The sample was quickly circulated in a closed circle-flow cell system with a total volume of 700μl for the fluorescence quenching experiment, and was illuminated with 90 mW of 370 nm light in one cuvette to prepare the *cis*-configuration, while the actual fluorescence experiment was performed in a second cuvette and induced with laser pulses at 266 nm with a pulse energy of 50 pJ at 2.5 kHz. During the course of an experiment (1 min), we estimated that only a tiny fraction of the molecules in the flow cell system ($2{\cdot}10^{-6}$) absorbed a 266 nm photon.

IV. COMPUTATIONAL DETAILS

All systems were solvated with ca. 7000-8000 TIP4P-2005[2] water molecules in a dodecahedron box with an image distance of 7-8 nm. Na^+ and Cl^- were added at a salt concentration of 0.1 M with an excess of Cl^- to compensate the net positive charge of the complexes. Force field parameters for the azobenzene switch were taken from Ref. 3. All bonds involving hydrogen atoms were constrained using the LINCS algorithm,[4] allowing for 2 fs time step. Long-range electrostatic interactions were computed by the Particle Mesh Ewald method,[5] whereas the short-range electrostatic interactions were treated explicitly with a Verlet cutoff scheme. The minimum cutoff distance for electrostatic and van der Waals interactions was set to 1.4 nm. A temperature of 300 K was maintained via the Bussi thermostat (aka velocity-rescale algorithm)[6] with a coupling time constant of 0.1 ps. A pressure of 1 bar was controlled using the pressure coupling method of Berendsen[7] with a coupling time constant of 0.1 ps. After energy minimization, all systems were equilibrated for 10 ns using an NPT ensemble. S-pep(6,15) was then simulated for 1 μs in both *cis* and *trans* configuration of photoswitch, S-pep(6,15)H12F for 800 ns in *cis* and 2 μs in *trans* configuration. S-pep(6,13), S-pep(6,10) and S-pep(7,11) were simulated for 400 ns in both configurations.

Principal component analysis on backbone dihedral angles (dPCA+)[8] was carried out by diagonalizing the covariance matrix $\sigma_{ij} = \langle(\varphi_i - \langle\varphi_i\rangle)(\varphi_j - \langle\varphi_j\rangle)\rangle$ of input coordinates $\{\varphi_i\}$. Projecting the MD trajectories onto the resulting eigenvectors $v^{(i)}$ we define the principal components, $V_i = v^i \cdot (\varphi - \langle\varphi_i\rangle)$, which are naturally ordered in descending variance of the system. After dimensionality reduction, robust density-based clustering[9] was performed by computing a local free energy estimate for every structure of the trajectory via counting all other structures inside a d-dimensional hypersphere of fixed radius R, where R equaled the lumping radius.[10] Normalization of these population count yields densities, which give the free energy estimate $F_R = -k_\mathrm{B}T\ln(P_R/P_R^{max})$ starting at min(F_R)= 0.

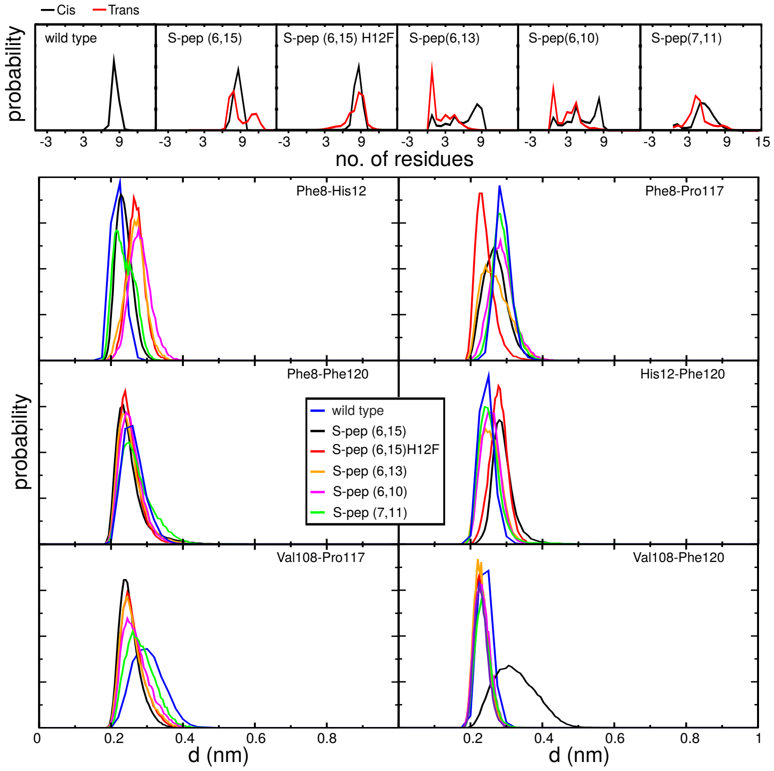
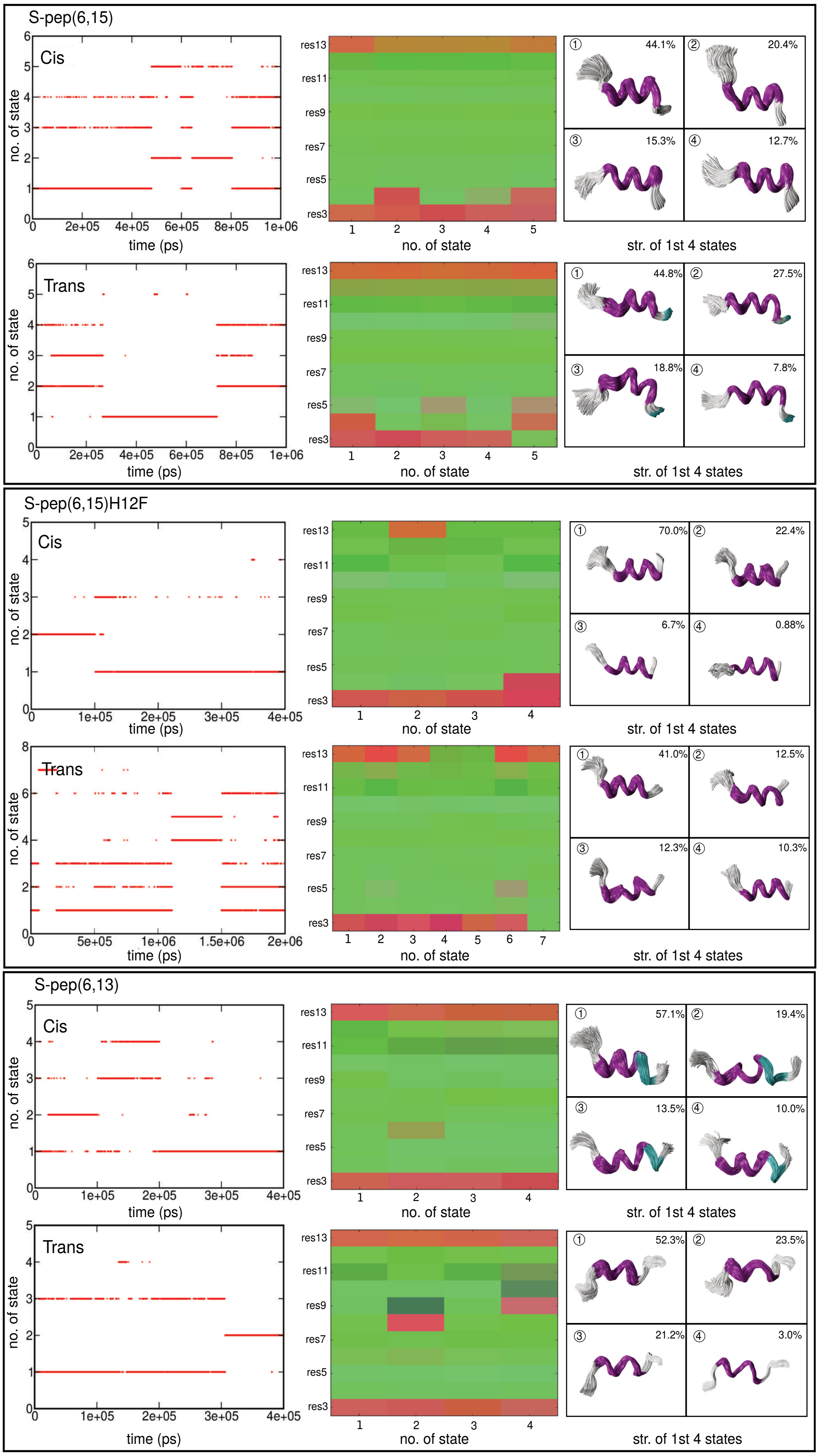


FIG. S3. Top: Probabilities of number of residues in a helical conformation; in black for the *cis*-state and in red for the *trans*-state. Bottom: Probability distribution of the most important contacts of *cis*-form peptides with S-protein. Note that the distributions are quite similar to results obtained for the wild-type system.

S5


S-pep(6,15)
Cis
Trans
no. of state
time (ps)
res13
res11
res9
res7
res5
res3
no. of state
str. of 1st 4 states
① 44.1%
② 20.4%
③ 15.3%
④ 12.7%
① 44.8%
② 27.5%
③ 18.8%
④ 7.8%
S-pep(6,15)H12F
① 70.0%
② 22.4%
③ 6.7%
④ 0.88%
① 41.0%
② 12.5%
③ 12.3%
④ 10.3%
S-pep(6,13)
① 57.1%
② 19.4%
③ 13.5%
④ 10.0%
① 52.3%
② 23.5%
③ 21.2%
④ 3.0%
0
2e+05
4e+05
6e+05
8e+05
1e+06
1e+05
3e+05
5e+05
1.5e+06
2e+06

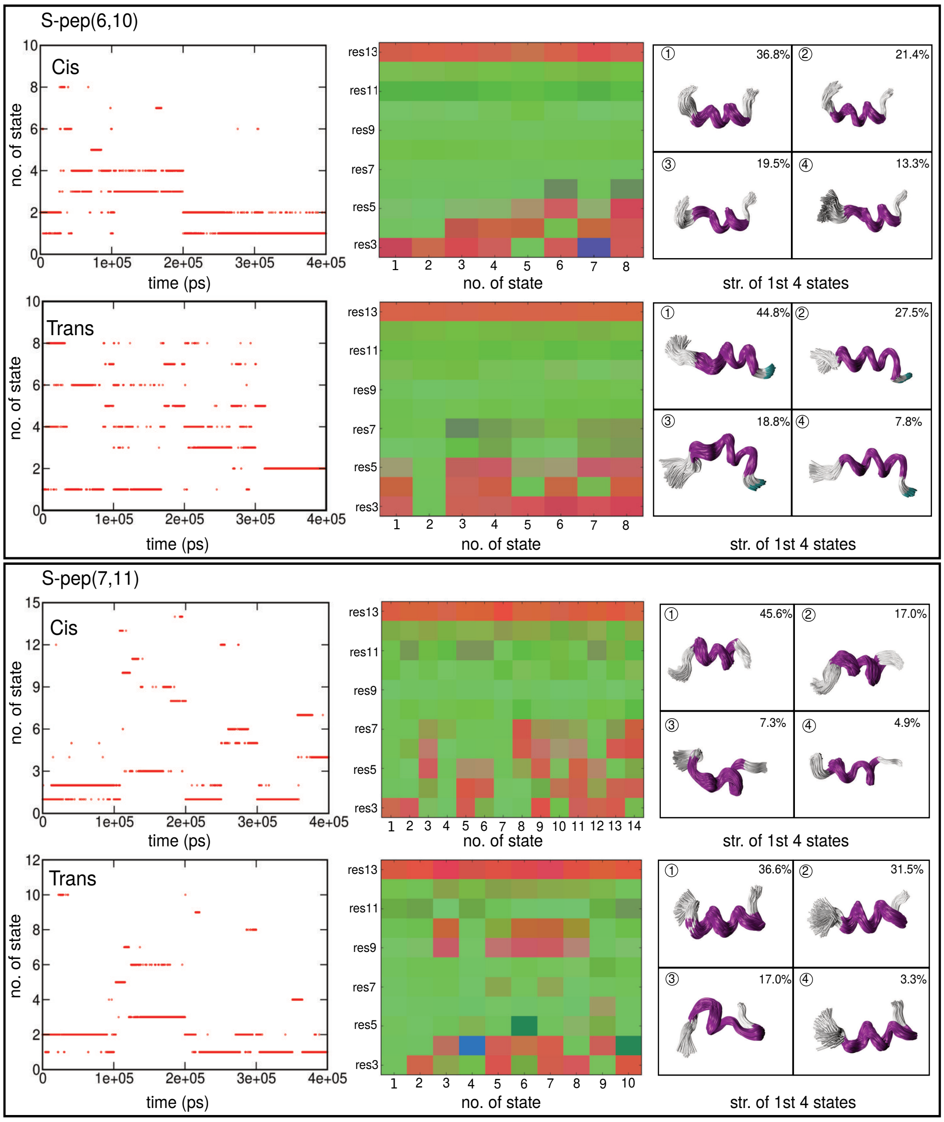


FIG. S4. State trajectory (No. of states vs No. of frames), ramacolor plot and representative structures of first four most populated states of all the considered systems in both *cis* and *trans*. The colors of the ramacolor plots are specified in Fig. S5.

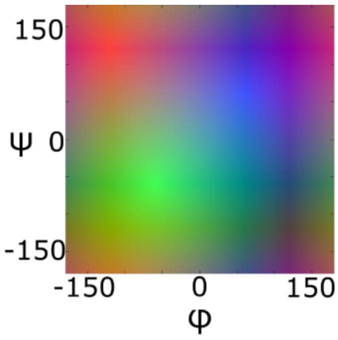


FIG. S5. Color space spanned by ϕ- and ψ-coordinates defined as: red is associated with extended conformations (β, P_{II}) centered around (-120, 120), green with right-handed helical conformations (α_R, 3_{10R} centered around (-60,-60), and blue with left-handed helical conformations (α_L, 3_{10L}) centered around (60, 60).

References:

(1) Dias, A. R.; Minas Da Piedade, M. E.; Martinho Simoes, J. A.; Simoni, J. A.; Teixeira, C.; Diogo, H. P.; Meng-Yan, Y.; Pilcher, G. Enthalpies of formation of cis-azobenzene and trans-azobenzene, *J. Chem. Thermodyn.* **1992**, *24*, 439–447.

(2) Abascal, J. L. F.; Sanz, E.; García Fernández, R.; Vega, C. A potential model for the study of ices and amorphous water: TIP4P/Ice, *J. Chem. Phys.* **2005**, *122*, 234511.

(3) Zanobini, C.; Bozovic, O.; Jankovic, B.; Koziol, K. L.; Johnson, P. J. M.; Hamm, P.; Gulzar, A.; Wolf, S.; Stock, G. Azidohomoalanine : A Minimally Invasive, Versatile, and Sensitive Infrared Label in Proteins To Study Ligand Binding, *J. Phys. Chem. B* **2018**, *122*, 10118–10112.

(4) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation, *Journal of Chemical Theory and Computation* **2008**, *4*, 435–447.

(5) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems, *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(6) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling, *J. Chem. Phys.* **2007**, *126*, 014101.

(7) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath, *J Chem Phys* **1984**, *81*, 3684–3690.

(8) Sittel, F.; Filk, T.; Stock, G. Principal component analysis on a torus: Theory and application to protein dynamics, *J. Chem. Phys.* **2017**, *147*, 244101.

(9) Sittel, F.; Stock, G. Robust density-based clustering to identify metastable conformational states of proteins, *J. Chem. Theory Comp.* **2016**, *12*, 2426–2435.

(10) Nagel, D.; Weber, A.; Lickert, B.; Stock, G. Dynamical coring of Markov state models, *J. Chem. Phys.* **2019**, page submitted.

5.1 S-pep(6,13)/S-protein Pump-Probe

$*$ I contributed to the work present in this section with the acquisition and analysis of the transient infrared data concerning the S-pep(6,13)/S-protein complex, which are shown in Fig. 5.1, 5.2, and 5.3. The design and realization of the S-pep(6,13)/S-protein system has been done by Dr. Brankica Janković.

The study of protein-peptide interactions has a crucial role in protein biochemistry and biophysics. Understanding the structural and dynamical aspects can give us insight into a multitude of biological regulation mechanisms. RNase S is a non-covalent enzyme complex, which is constituted of two subunits: S-peptide and S-protein [170]. It has to be stressed that neither S-peptide nor S-protein are enzymatically active when one of the two binding partners is missing. Moreover, S-peptide adopts a disordered and unfolded structure when it is unbound from S-protein, while S-protein assumes a much more flexible conformation when S-peptide is absent [74, 171–174]. These characteristics make RNase S a prototype model system to be investigated in order to elucidate the binding mechanism between S-protein and S-peptide, whose comprehension is still missing [175–177]. Two limiting binding scenarios named "induced fit" and "conformational selection" have been proposed [178] to describe the S-peptide and S-protein binding mechanism, where the difference between the two depends on the chronological order of the binding and conformational change events, *i.e.* kinetics of binding versus conformational change. Probing the molecular vibrations of the S-peptide/S-protein adduct *via* transient infrared (IR) spectroscopy can shed light on the binding mechanism involved in the formation of the RNase S complex, providing us with information about the kinetics involved in the (un)binding process.

In this section, some of the time-resolved data in regards to the RNase S complex are reported. Among the five photoswitchable S-peptide mutants proposed by Janković *et al.* [148], S-pep(6,13) is the most interesting variant, due to its capability to switch on and off the binding affinity towards S-protein, depending on the isomerization state of the photoswitch. The data were collected as a visible pump - infrared probe type of experiment (see Methods for details – section A.2 – Transient Infrared (IR) Spectroscopy), where the visible pump pulse ($\lambda_{pump} \approx 420$ nm, energy per pulse of ≈ 3 μJ, pulse duration ≈ 5 ps) induced the *cis* to *trans* photoisomerization of S-pep(6,13), while the infrared probe pulse monitored the structural rearrangements of the overall S-pep(6,13)/S-protein system, tracking the transient amide I response of it.

The transient infrared data (see Fig. 5.1) reveal several absorption bands in the spectral window from ≈ 1560 cm^{-1} to ≈ 1700 cm^{-1}. As described in Methods (section A.3 – Lifetime Analysis), the maximum entropy inversion of the Laplace

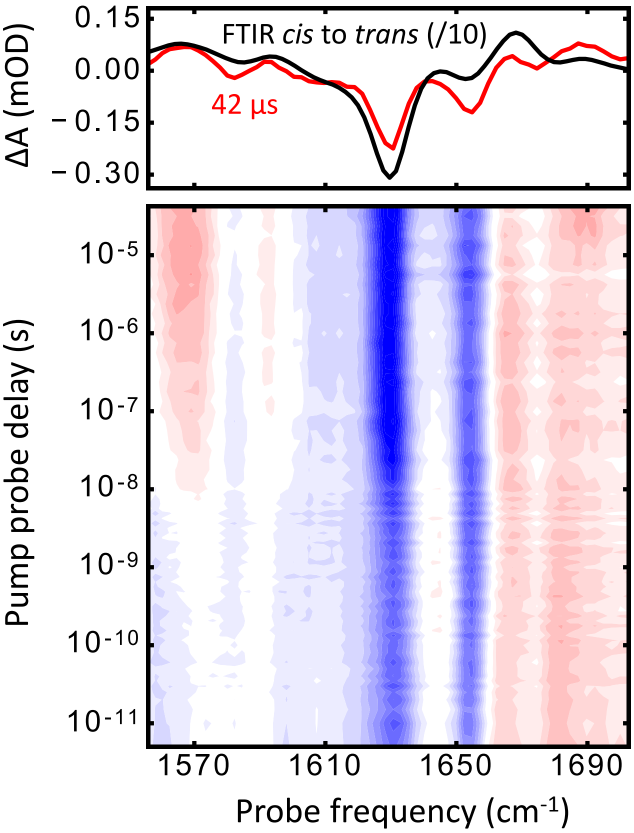


Figure 5.1: *Cis* to *trans* S-pep(6,13)/S-protein pump-probe data. The transient infrared spectra have been collected from ≈ 5-10 ps until ≈ 42 μs. The top panel shows the comparison between the transient infrared spectrum acquired at the longest pump-probe delay time of ≈ 42 μs (red line) with the *cis* to *trans* FTIR difference spectrum (black line and properly scaled) of the same sample, which represents *de facto* the spectrum acquired at infinite time after photoswitching.

transform [200, 201] has been employed to analyse the transient infrared data, obtaining the lifetime spectrum in panel (a) of Fig. 5.2. The latter graph provides us with indications about the complexity of the overall dynamics of the S-pep(6,13)/S-protein system upon *cis* to *trans* peptide photoswitching. In order to have a better overview of it, the averaged lifetime spectrum $a(\tau_j)$ in panel (b) of Fig. 5.2 has been calculated, giving us a direct understanding of the timescales involved in the overall non-equilibrium process. Several peaks have been observed, which are spread over the whole time window accessible with the pump-probe experiment, starting from a few ps up to *circa* 42 μs, which have been highlighted with horizontal black dotted lines. In details, the seven peaks have the maximum at ≈ 13 ps, 130 ps, 1.3 ns, 20 ns, 250 ns, 2.5 μs and 20 μs, where the one at ≈ 20 ns is the most intense among all of them, and it is mainly due to the band at ≈ 1628 cm^{-1} (see Fig. 5.3). Moreover, from the comparison of the transient infrared spectrum acquired at the longest pump-probe delay time of ≈ 42 μs (red line) with the properly scaled *cis* to *trans*

FTIR difference spectrum (black line) in the top graph of Fig. 5.1, we can conclude that the S-pep(6,13)/S-protein system undergoes significant structural rearrangements within 42 μs. However, from the discrepancy between the red and black lines, it seems that the overall non-equilibrium process is not completely finished within this time scale.

In conclusion, the perturbation of the RNase S adduct upon *cis* to *trans* S-pep(6,13) photoswitching reveals a discrete non-equilibrium dynamics of the system, where several maxima appear from the lifetime analysis in panel (b) of Fig. 5.2, highlighting in this way the time scale separation of the overall non-equilibrium process from $\approx$ 5-10 ps until $\approx$ 42 μs. This discrete set of time scales has been highlighted employing horizontal black dotted lines in Fig. 5.2, connecting the maxima observed in the average lifetime spectrum in panel (b) with the lifetime spectrum in panel (a). It has to be stressed that at this stage, the non-equilibrium dynamics of the system

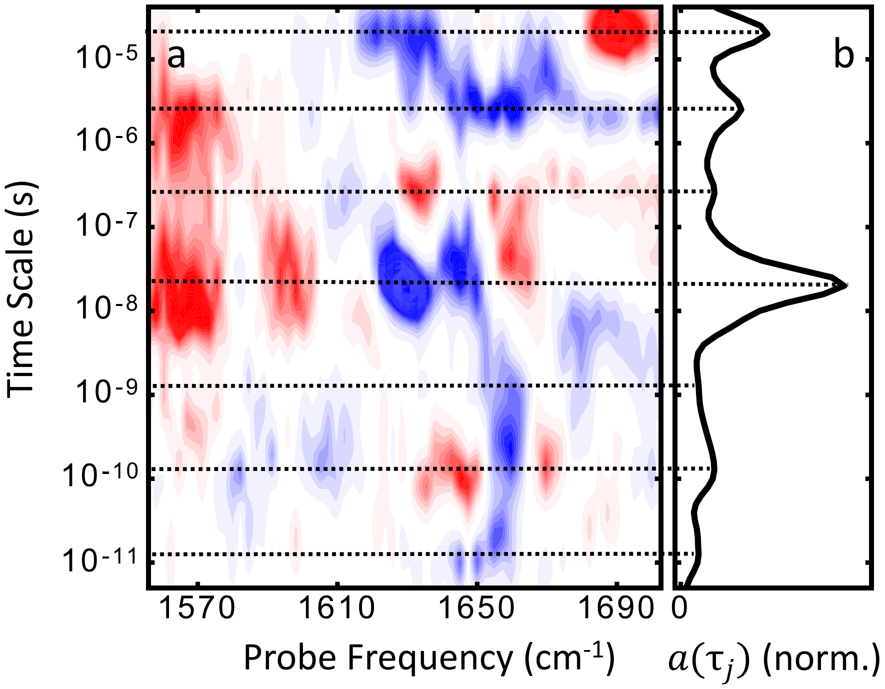


Figure 5.2: *Cis* to *trans* S-pep(6,13)/S-protein lifetime analysis. (a) Lifetime spectrum $a(\omega_i, \tau_j)$ (see Methods for details, section A.3 – Lifetime Analysis – Eq. A.1) of the *cis* to *trans* S-pep(6,13)/S-protein pump-probe data in Fig. 5.1. Red colours indicate positive amplitudes $a(\omega_i, \tau_j)$, while blue colours negative amplitudes $a(\omega_i, \tau_j)$. (b) Averaged lifetime spectrum $a(\tau_j)$; the 2D information present in panel (a) has been averaged in an absolute sense over the probe frequency axis ω_i and squeezed in a 1D plot (see Eq. A.5). The horizontal black dotted lines connecting the lifetime spectrum $a(\omega_i, \tau_j)$ in panel (a) with the corresponding average lifetime spectrum $a(\tau_j)$ in panel (b) highlight the discretization of time scales in regards to the overall non-equilibrium dynamics of the S-pep(6,13)/S-protein system upon *cis* to *trans* S-pep(6,13) photoswitching.

observed in Fig. 5.2 can be assigned neither to S-pep(6,13) nor to S-protein, due to the highly convoluted nature of the overall amide I response of the RNase S adduct, which cannot be easily disentangled. In order to separate and assign the various contributions, isotope labeling strategy is needed [46–50] as employed in the PDZ2 domain case (see chapter 4 for details). Furthermore, it would be helpful to know the timescale of the *cis* to *trans* unbinding event of S-pep(6,13) from S-protein in order to understand if the structural rearrangements observed up to $\approx 42\ \mu$s (see Fig. 5.1 and Fig. 5.2) can be correlated or not with it. All these experiments are underway.

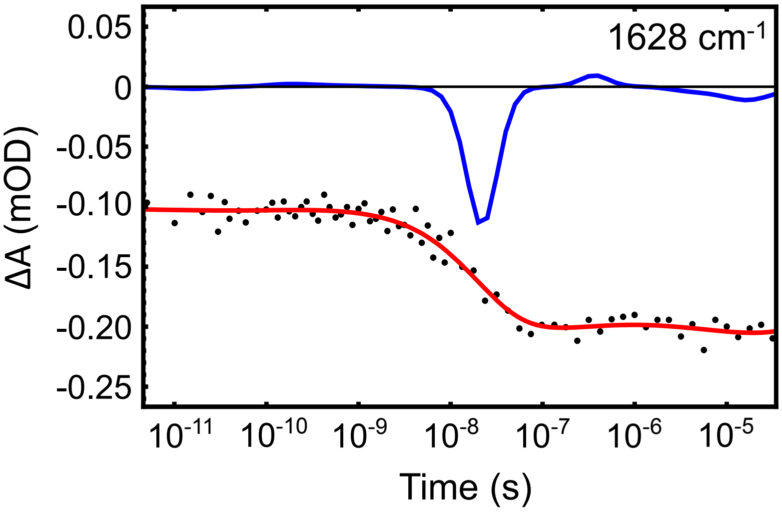


Figure 5.3: S-pep(6,13)/S-protein transient kinetic trace at ≈ 1628 cm^{-1}. The experimental data are shown as black dots, the fit $S(\omega_i, t)$ (see Methods for details, section A.3 – Lifetime Analysis – Eq. A.1) as red lines, and the amplitudes $a(\omega_i, \tau_j)$ of the lifetime spectrum in Fig. 5.2 at this probe frequency as blue lines. It is evident from the experimental data that the main change of the transient infrared signal at ≈ 1628 cm^{-1} happens on a timescale of ≈ 20 ns, where the amplitudes $a(\omega_i, \tau_j)$ are the biggest. This band contributes to a great extent to the overall non-equiibrium dynamics of the S-pep(6,13)/S-protein complex upon *cis* to *trans* S-pep(6,13) photoswitching, as highlighted in the averaged lifetime spectrum $a(\tau_j)$ (see Fig. 5.2, panel (b)).

Chapter 6

Conclusions and Outlook

Molecular spectroscopy, *i.e.* the study of light-matter interactions, allows us to excite specific transitions in a selective manner, which can be *e.g.* vibrational or electronic transitions if an IR or UV-VIS pulse is employed, respectively. This approach enables us to gain information about the molecular system under investigation from specific perspectives, like in the cases of the azidohomoalanine (Aha) mutants and of the non-equilibrium protein response upon peptide photoswitching concerning the PDZ2 domain system, as described in chapters 2, 3, and 4 of this thesis. Regarding the azobenzene photoswitch involved in both PDZ2 domain and RNase S complex, the use of a UV-VIS pulse allows the system to overcome the activation energies (E_a) for the *cis* to *trans* and *trans* to *cis* isomerization processes, which are *circa* 30 and 50 times bigger than the thermal energy available at room temperature, respectively ($E_{a(cis \rightarrow trans)} \approx$ 74-78 kJ/mol [141], $E_{a(trans \rightarrow cis)} \approx$ 125 kJ/mol, $k_B T_{298K} \approx$ 2.5 kJ/mol). The photoinduced isomerization events for both photoswitching directions are fully reversible and happen on a ps timescale, allowing reactions otherwise inaccessible by thermal processes, which can be followed in an ultrafast time-resolved manner.

In this thesis, the study of a PDZ2 domain system has been presented, where the site-specific infrared label azidohomoalanine [35, 36, 99–103] has been exploited as a local infrared reporter for the PDZ2 domain chemical environment by the use of 2D-IR spectroscopy [35, 101], and a non-equilibrium allosteric transition of the PDZ2 domain has been investigated, employing TRIR spectroscopy together with isotope labeling strategy [46–50]. As a starting point, Aha has been inserted in six different positions (one at the time) within the protein structure to verify if the structural changes between the apo and holo states of the PDZ2 domain (with ligand unbound and bound, respectively) could have been sensed by Aha to obtain site-specific information about the overall apo/holo process (see chapter 2 for details) [101]. The ligand binding process has been mimicked (it has been confirmed by NMR spectroscopy) by photoisomerization of an azobenzene moiety covalently linked to the binding groove of the PDZ2 domain, where the *cis* and *trans* configurations of the photoswitch [131, 141, 143] mimicked the apo and holo states of the PDZ2 domain [145]. Interestingly, one of the six Aha mutants showed a remarkable 2D-IR difference signal upon photoswitching, where a surprising and uncommon result has been observed, namely an intensity change of the Aha absorption band without any significant frequency shift of it. Subsequently, another Aha-mutated version of the same PDZ2 domain without any photoswitch present within its structure (the K38Aha mutant) has been exploited to investigate the protein-ligand recognition event in a minimally invasive manner, where two different peptide-ligands have been employed, observing small but reproducible Aha vibrational shifts of *circa* 1-3 cm^{-1} upon protein-peptide binding (see chapter 3 for details) [35]. One of the two peptide versions has been designed in such a way that its binding affinity towards the PDZ2 domain changes upon photoisomerization of an azobenzene moiety covalently linked to it. The experimental evidence that the ligand recognition event or the change of the ligand binding affinity (*e.g.* by phosphorylation) in PDZ domains influences the side-chain fluctuation dynamics of the protein, which is known to be related with the allosteric response of PDZ domains [59–65], suggests an excellent candidate to study the PDZ2 domain structural rearrangements associated to the ligand binding process in an ultrafast time-resolved manner. Hence, the photoswitchable peptide has been employed as a phototrigger to promote a non-equilibrium allosteric transition of the PDZ2 domain as a consequence of the different binding affinities of the ligand in the *cis* and *trans* states (K_D difference of *circa* 5 fold). Structural fluctuation dynamics of several PDZ domains have been already investigated by NMR spectroscopy [59–61, 63–65, 205], however with regards to equilibrium relaxation dynamics [66, 67] and not to non-equilibrium dynamics, where an intrinsically higher time resolution, which is provided by the use of IR spectroscopy, is needed [145, 146]. In order to isolate the TRIR response of the PDZ2 domain from the one of the photoswitchable peptide, an isotope labeled ($^{13}\mathrm{C}^{15}\mathrm{N}$) PDZ2 domain version has been expressed and employed for this purpose. The protein response has been investigated under two photoswitching approaches, where the *cis* to *trans* and *trans* to *cis* light-induced peptide isomerizations have been promoted (see chapter 4 for details). In short, two almost identical samples have been used, which differ only for the presence of a heavier ($^{13}\mathrm{C}^{15}\mathrm{N}$) PDZ2 domain isotopologue in one case, while the photoswitchable peptide has the naturally abundant ($^{12}\mathrm{C}^{14}\mathrm{N}$) isotopic composition in both cases, and it is bound to the protein. As a result of the isotope labeling effect [46–50], the amide I band of the protein has been red-shifted, while the amide I band of the photoswitchable peptide has not been altered, allowing us for the isolation of the PDZ2 domain response upon peptide photoswitching. It has to be stressed that this perturbative approach does not involve any chemical modification of the PDZ2 domain, but rather it concerns only the ligand, which is responsible for promoting the allosteric transition of the PDZ2 domain. From the transient infrared data, it emerges that the non-equilibrium response of the PDZ2 domain is not symmetric, namely it follows different evolution pathways when the *trans* to *cis* and *cis* to *trans* photoisomerizations are promoted. Moreover from the lifetime analysis, a small number of discrete time scales have been observed for both photoswitching approaches, emphasizing the time scale separation involved in the overall non-equilibrium allosteric transition of the PDZ2 domain. These experimental observations can be interpreted with a Markov state model [183–188], where the relative populations of the Markov states, which represent the PDZ2 domain structural ensemble, are different in the two cases.

In parallel, another biological system has been investigated with the intent to modulate by the light reversibly, the binding affinity between the two subunits (S-protein and S-peptide) which constitute the adduct ribonuclease S (RNase S) [74, 75, 175, 176, 179, 202]. Five photoswitchable S-peptide variants have been designed and synthesized, where the binding affinity modulation has a different extent in each case, regarding the *trans* or *cis* isomerization state of the photoswitch. In the specific case of S-pep(6,13), this binding affinity modulation is maximal; in the *cis* configuration, S-pep(6,13) has a dissociation constant (K_D) value of $\approx$ 70-130 μM, while for the *trans* form, non-specific binding has been observed (see chapter 5 for details) [148, 203, 204]. The experimental K_D values obtained for the five photoswitchable S-peptide variants correlate well with their amount of α-helicity, as predicted by molecular dynamics (MD) simulations. Moreover, the S-protein/S-pep(6,13) model system opens new scenarios towards novel time-resolved unbinding experiments, where a big K_D value is needed to speed up the unbinding event. For example, considering the values for the wild type S-peptide from the paper of Bachmann *et al.* [175], with a K_D value of 0.28 μM and a kinetic association constant (k_{on}) value of 4.4×10^5 M^{-1}s^{-1}, the calculated kinetic dissociation constant (k_{off}) is 0.12 s^{-1} with a half-life ($t_{1/2}$) for a first order kinetic reaction of 5.7 s. In principle, considering a hypothetical K_D value of 1 mM and assuming the same k_{on} value as before, the calculated k_{off} is 4.4×10^2 s^{-1} with a $t_{1/2}$ of 1.6 ms, which results in a faster unbinding process than in the previous case. While a reasonable k_{off} estimate can be calculated for the *cis* configuration of S-pep(6,13), once both K_D and k_{on} for the *cis* state are known ($k_{off} = K_D * k_{on}$), the same is not doable for the *trans* isomerization state of S-pep(6,13). The reason behind it is that S-pep(6,13) binds in a specific manner, *i.e.* it binds to the binding groove of S-protein only in the *cis* configuration, while in the *trans* isomerization state the binding is non-specific [148, 203, 204], namely it binds to a multitude of distinct sites of S-protein and neither K_D nor k_{on} values are representative of the S-pep(6,13) binding event on the binding groove of S-protein. Nevertheless, one way to determine the k_{off} value for the *trans* isomerization state of S-pep(6,13) from the binding pocket of S-protein is *via* transient fluorescence spectroscopy, preparing the RNase S complex with S-pep(6,13) in the *cis* isomerization state, and instantaneously switch the latter to the *trans* configuration, employing an actinic pump pulse. In this way, the time scale for the desired unbinding event could be obtained experimentally, due to the different fluorescence quenching in the *cis* and *trans* cases, where the overall fluorescence emission of the six tyrosine (Y) residues of S-protein is affected by the proximity of S-pep(6,13) to it [148, 206, 207]. Once the experimental k_{off} is known, the unbinding event could be detected from the perspective of the structural changes associated to it for both constituents of the RNase S adduct, S-protein and S-pep(6,13). The transient infrared data, which have been shown in section 5.1, represent the overall amide I non-equilibrium response of the naturally abundant ($^{12}C^{14}N$) S-protein/S-pep(6,13) complex upon *cis* to *trans* S-pep(6,13) photoswitching. From the experimental data and the lifetime analysis, it emerges that, also for the RNase S complex as well as for the PDZ2 domain case, a discrete number of time scales are characteristic of the non-equilibrium response of the system under investigation, where the most significant structural changes of it happen on a time scale of *circa* 20 ns. The combination of TRIR spectroscopy with isotope labeling strategy [46–49] could clarify the structural rearrangements of the system before and after the unbinding event by a comparison among time scales, elucidating in this way the binding mechanism involved in the formation/dissociation of the RNase S complex, namely if it can be classified as an induced fit or a conformational selection type of mechanism [178]. Having available both naturally abundant ($^{12}C^{14}N$) and heavier ($^{13}C^{15}N$) isotopologues for either S-protein and S-pep(6,13) variant, the TRIR response in the amide I spectral region could be disentangled for both subunits, isolating in this way the responses of S-protein and S-pep(6,13) in the RNase S complex upon S-pep(6,13) photoswitching.

In addition to this challenging outlook proposed for the RNase S complex, another model system is now presented: a photocontrollable PDZ3 domain. From a structural point of view, the main difference with the previously described PDZ2 domain is the presence of an additional third α-helix ($\alpha3$) in the structure of the PDZ3 domain, which lies far away from the binding groove of the protein [65, 195, 208–211]. This PDZ3 domain is part of the multidomain complex postsynaptic density protein 95 (PSD-95), which is a member of the membrane-associated guanylate kinase (MAGUKs) family, *i.e.* a class of multidomain proteins of paramount importance in science for their involvement in cell signaling [57, 86]. There are several lines of evidence showing that the modification of this extra $\alpha3$-helix, by phosphorylation or by its removal, significantly alters the binding affinity of the PDZ3 domain towards the recognition sequence of the C-terminal peptide-ligand, which binding to the protein induces a dynamic allosteric effect in the PDZ3 domain [208, 211]. The fascinating idea is to attach the photoswitch to the $\alpha 3$-helix, in order to photocontrol the α-helical folding/unfolding of it reversibly, and to observe how the ligand binding event is affected by photoswitching. The presence of an allosteric network between the $\alpha 3$-helix and the binding groove of the PDZ3 domain should reflect in a change of the peptide-ligand binding affinity. However, if the effect of photoswitching the $\alpha 3$-helix on the change of the binding affinity is tiny and not easily detectable employing common experimental methods for K_D determination like isothermal titration calorimetry (ITC), tryptophan (W) and/or tyrosine (Y) fluorescence quenching, *etc.* [148], the allosteric effect could be observed from the point of view of the photoswitch, monitoring the changes of its *cis* to *trans* thermal relaxation kinetics upon ligand binding. If there is a correlation between the structure of the $\alpha 3$-helix and its fluctuation dynamics with the binding of the ligand, the photoswitch should sense a different strain upon ligand binding, which would affect its thermal relaxation kinetics. It has to be stressed that the latter approach is more sensitive than employing binding affinity determination methods, due to the exponential dependence of the thermal relaxation kinetic from the activation energy (E_a). In other words, tiny changes of E_a provoke significant changes in the thermal relaxation kinetic of the photoswitch, which can be easily observed experimentally. Moreover, TRIR spectroscopy experiments could be performed, triggering the $\alpha 3$-helix disruption or folding in order to investigate the protein structural rearrangements associated with it, in an ultrafast time-resolved manner. Hence, the above experiments could shed light on the allosteric network between the $\alpha 3$-helix and the binding groove of the PDZ3 domain, getting insight into the time scales involved in the overall allosteric process. All these experiments are on their way.

Appendix A

Methods

A.1 Determination of the Binding Affinities

The isothermal titration calorimetry (ITC) measurements were performed using a MicroCal ITC200 (Malvern, UK), with the same stock solutions of the photoswitchable peptide-ligand and of the PDZ2 domain for the *trans* and *cis* ITC experiments, and under exactly the same experimental conditions, in order to obtain mutually comparable dissociation constants (K_D) values. The experiments were done in triplicate to ensure the reproducibility of the data. The sample cell was loaded with 250 μl of 80 μM PDZ2 domain solution and the syringe was loaded with 40 μl of 800 μM photoswitchable peptide-ligand solution. For the *trans* measurement, the system was kept in the dark for the duration of the whole experiment, while for the *cis* measurement the syringe was constantly illuminated employing a 370 nm cw laser (CrystaLaser, power $\approx$ 90 mW) [148], obtaining the thermograms in Fig. 4.2. In addition, circular dichroism (CD) spectroscopy and intrinsic tryptophan (W) fluorescence quenching were used as alternative methods to determine the binding affinities of the *trans* and *cis* states. Both spectroscopic signals showed a change upon the formation of the protein-ligand complex, which was proportional to the amount of complex formed in solution. This signal change was used to obtain the binding affinity curves, which were fit assuming a bimolecular equilibrium to obtain the dissociation constants (K_D) values. Circular dichroism (CD) measurements were performed using a Jasco (Easton, MD) model J810 spectropolarimeter in a 0.1 cm quartz cuvette and intrinsic tryptophan (W) fluorescence quenching employing a PelkinElmer spectrofluorimeter as described by Janković *et al.* [148]. In either case, the protein concentration was kept constant at 5 μM, while the peptide concentration was varied. The results obtained from circular dichroism (CD) and intrinsic tryptophan (W) fluorescence quenching spectroscopies are showed in Fig. 4.3, while Tab. 4.1 compares the dissociation constants (K_D) values obtained from all different methods.

A.2 Transient Infrared (IR) Spectroscopy

The transient infrared (IR) spectra reported in Fig. 4.4 and 5.1, were collected as a VIS-pump-IR-probe type of experiment, employing two electronically synchronized femtosecond Ti:Sapphire laser systems [212] running at 2.5 kHz. Mid-IR probe pulses centered at ≈ 1630 cm^{-1}, with a duration of ≈ 100 fs and a beam diameter on the sample of ≈ 150 μm, were obtained employing a home-made optical parametric amplifier (OPA) [213]. The probe pulses were dispersed through a spectrograph and detected employing a 2×64 MCT array detector with a spectral resolution of ≈ 2 cm^{-1} per pixel. The pump pulses had a beam diameter on the sample of ≈ 180 μm and their wavelength was tuned to obtain 420 and 380 nm radiation *via* second harmonic generation in a BBO crystal to promote the *cis* to *trans* and *trans* to *cis* photoisomerizations of the photoswitchable peptide-ligands, respectively. It has to be stressed that the experimental conditions employed to collect the transient IR spectra in Fig. 4.4 and 5.1 were different. In the case of the PDZ2 domain in Fig. 4.4, the pump pulses were stretched to *circa* 200 ps, bypassing the compression stage in the regenerative amplifier to minimize the sample degradation during the transient infrared experiments. The energies per pulse employed in this case were *circa* 2.1 and 1.3 μJ for the *trans* to *cis* and *cis* to *trans* photoswitching experiments, respectively. It has to be underlined that to acquire the transient IR spectra, the peak power of the pump pulses was decreased considerably, avoiding in this way aggregation and disruption of the sample at the expense of a lower experimental time resolution. The sample concentrations were set ≈ 1.25 mM for the photoswitchable peptide-ligand and ≈ 1.5 mM for the PDZ2 domain. In these conditions, together with the K_D values of ≈ 2.0 μM and ≈ 9.6 μM for the *trans* and *cis* states, the photoswitchable peptide-ligand is essentially bound to the PDZ2 domain in both *trans* and *cis* configurations (fraction bound $> 95\%$ in both cases). The samples were dialyzed against 50 mM borate, 150 mM NaCl buffer at pH = 8.5, lyophilized, then resuspended and incubated in D_2O overnight at room temperature, lyophilized and resuspended in D_2O again before collecting the transient IR measurements, eliminating in this way the H/D exchange during the experiments. The pump-probe spectra were acquired up to the maximum delay time of ≈ 42 μs with a time resolution of ≈ 200 ps, employing the normalisation for the noise suppression as described by Feng *et .al* [214]. In the case of the S-pep(6,13)/S-protein complex in Fig. 5.1 the transient IR data were collected from $\approx$ 5-10 ps until ≈ 42 μs pump-probe delay times with a time resolution of ≈ 5 ps in this case. The pump pulses (λ_{pump} ≈ 420 nm) were stretched to ≈ 5 ps to minimize the sample degradation during the *cis* to *trans* S-pep(6,13) photoswitching experiment, employing an energy per pulse of ≈ 3 μJ. The concentrations of S-protein and of S-pep(6,13) in the sample were set $\approx$ 1.25 mM and $\approx$ 1.5 mM, respectively. In these conditions, the amount of S-pep(6,13) bound to S-protein in the *cis* configuration is $\approx$ 72 %, taking into account an experimental K_D value of $\approx$ 70 μM for the *cis* configuration, while it is not determined for the *trans* isomerization state, due to the non-specific binding of S-pep(6,13) in this case [148, 203, 204]. The buffer used for the S-pep(6,13)/S-protein complex was 50 mM sodium phosphate at pH = 7.0, employing the same procedure as before to eliminate the H/D exchange during the transient IR experiments. All the samples were circulated ($\approx$ 700 μL) using a peristaltic pump purged with N_2, a reservoir, and a flow-cell mounted with a Teflon spacer of $\approx$ 50 μm thickness at a flow rate of $\approx$ 1 mL per minute. The flow speed in the sample cell was optimized to minimize the loss of sample at the longest pump-probe delay time ($\approx$ 42 μs) and to exchange the sample completely for the subsequent laser shot after 400 μs at the same time.

For the *trans* to *cis* photoswitching experiment, the thermal *cis* to *trans* relaxation was exploited to ensure that the transient IR signal was representative of the *trans* to *cis* photoswiching direction. In other words, comparing the thermal *cis* to *trans* relaxation rate (half-life ($t_{1/2}$) *circa* 19 min for the photoswitchable peptide-ligand in the PDZ2 domain case) with the isomerization probability induced by the 380 nm pump light (given by light power, total sample volume, molar extinction coefficient [131], and isomerization quantum yield [141]), the estimate of the photoequilibrium position in the total sample volume was 70%/30% *trans*/*cis* during the measurement. Moreover, the absorption cross section at 380 nm of the azobenzene moiety in the *trans* configuration is $\approx$ 20 times bigger than that of the *cis* isomer [131], meaning that $>$ 97% of the molecules in the *trans* to *cis* experiment undergo the desired isomerisation direction. For the *cis* to *trans* photoswitching experiments, the samples were illuminated with an excess of light in a reservoir using a cw laser ($\lambda_{emission} \approx$ 370 nm, power $\approx$ 150 mW, CrystaLaser) in order to promote the inverse *trans* to *cis* photoisomerization, estimating in this case that $>$ 80% of the transient IR signal originated from the *cis* to *trans* photoisomerization process.

A.3 Lifetime Analysis

In essence, the transient infrared data $S(\omega_i, t)$ in Fig. 4.4 and Fig. 5.1 were fitted in the time domain (t) with the following multiexponential function:

$$S(\omega_i, t) = a_0(\omega_i) - \sum_j a(\omega_i, \tau_j) e^{-t/\tau_j} \tag{A.1}$$

where the index i refers to the kinetic trace at the probe frequency ω_i, and j to the time constant τ_j. The time constants were fixed and distributed on a logarithmic scale with ten terms per decade, while the amplitudes $a(\omega_i, \tau_j)$ were set as the free fitting parameters. Inversion of Eq. A.1 is an ill-posed problem, in other words, let y and x be L^2 space functions, and K a continuous linear operator between them, where:

$$y = Kx \tag{A.2}$$

The direct problem is to compute y given x, while the inverse problem is to compute x given the experimental data y. The inverse problem is well-posed when the solution exists, it is unique, and it is stable (it depends continuously from the data y) [215]. If one of the above conditions is not satisfied, the problem is ill-posed [216, 217]. Therefore, it is necessary to incorporate additional information about the desired solution to stabilize the problem and to avoid ill-posedness; regularization procedures have been introduced to single out a useful and stable solution [218]. Most probably, the best-known regularization mechanism is the one of Tikhonov [219], which selects the solution x_λ which solves the minimization problem

$$\min_x \{ \, \|(Kx - y)\|_2^2 + \lambda^2 \, \|x\|_2^2 \, \} \tag{A.3}$$

where $\| . \|_2$ is the Euclidean norm, and λ is a parameter which controls the relative weight of $\|x\|_2$ on the residual $\|(Kx - y)\|_2$ in the overall minimization procedure. The regularization term, $\|x\|_2$ in the Tikhonov case, is referred sometimes as the "smoothing term" or as the "penalty function", which has the role in adding constraints to the minimization procedure in order to obtain a stable solution as close as possible to the most representative one from the experimental data.

The maximum entropy method (MEM) [220, 221] is another regularization technique which has the advantage to allow us to obtain information with the least biased estimation possible from the experimental data [222, 223]. In other words, the penalty function in the MEM is the Shannon entropy [224, 225] $S(p) = -\sum p \log p$, where the chosen solution x_λ is the one which solves the minimization problem

$$\min_x \{ \, \|(Kx - y)\|_2^2 + 2\lambda \sum_i x_i \log x_i \} \tag{A.4}$$

The choice of the Shannon entropy as the penalty function and the logarithmic nature of it implies that the solution must have non-negative values and that the latter has not a linear dependency with the observation vector y. The positivity and non-linearity of the MEM confer benefits in several applications like astronomy [226], NMR [227] and time-resolved spectroscopy [200, 201, 228], which are worth the computational cost of the MEM. The algorithm employed to obtain the lifetime spectra in panels (a, c) of Fig. 4.6, and in panel (a) of Fig. 5.2 took into account the possibility to have both positive and negative amplitudes $a(\omega_i, \tau_j)$ in Eq. A.1, see Ref. [200] and Ref. [229] for details. The averaged lifetime spectra $a(\tau_j)$ in panels (b, d) of Fig. 4.6, and in panel (b) of Fig. 5.2 have been calculated in the following manner:

$$a(\tau_j) = \frac{\sqrt{\sum_{i=1}^{n} a(\omega_i, \tau_j)^2}}{n} \tag{A.5}$$

where n represents the total number of probe frequencies ω_i which are present in the corresponding lifetime spectra, and $a(\omega_i, \tau_j)$ are the free fitting parameter in Eq. A.1 employed with the maximum entropy method (MEM).

A.4 The 2D-IR Spectrometer in the Box-CARS Geometry

∗ I contributed to the work present in this section with the improvement of the existing 2D-IR spectrometer in the box-CARS geometry (see Fig. A.1) [31, 35, 101–103, 230] by exact-quasi-phase-cycling scattering suppression as described by Bloem *et al.* [231] implemented at the repetition rate of ≈ 5.13 kHz, and by polarization control of the pump excitation [232], designing and employing the "special mask" described in Fig. A.2. Moreover, a novel method has been proposed to suppress to a great extent the D_2O absorption background from the azidohomoalanine (Aha) absorption signal (see subsection A.4.1 – Pump-Induced Linear Dichroism Spectra, Fig. A.4, Fig. A.7 and Tab. A.2 for details), which is possible using a static cell with only ≈ 0.5 μL of ≈ 1.2 mM of Aha sample solution, without acquiring and subtracting the 2D-IR spectrum of the buffer alone.

Despite the complexity of the 2D-IR spectrometer in the box-CARS geometry and its specificity towards azidohomoalanine (Aha) related issues, several PhD students and Post Docs including Robbert Bloem, Klemens Lucas Koziol, Dr. Philip J. M. Johnson and I spent time and effort for its realization and implementation. Here, its latest version to date (May 2019) has been presented in detail as a hint for people who would like to use it thoroughly and understand its full potential.

As a source of mid-IR pulses, the output from a home-built optical parametric amplifier (OPA) [213] was used (pulse duration FWHM ≈ 100 fs, energy per pulse ≈ 3 μJ, pulse spectrum centered at ≈ 4.74 μm, laser repetition rate ≈ 5.13 kHz). A curved mirror (f = 315 mm) was employed to collimate the outgoing p-polarized mid-IR beam from the OPA, which was then spatially overlapped with a s-polarized continuous wave (cw) HeNe laser ($\lambda = 632.8$ nm, 05-LHP-151, Melles Griot, USA) by the use of a germanium plate set at the Brewster angle [233]. The co-propagating beams were split into three equally intense parts by custom-made BaF_2 beam splitters (Laseroptik) to generate the beams 1, 2, and 3. A small reflection of beam 3 taken from a CaF_2 window was used to generate the local oscillator (LO), and a small fraction of it was employed to monitor the quality of the mid-IR source with the help of a single-pixel MCT detector (Infrared Associates, USA) while measuring the 2D-IR spectra. Two voice coil stages (Equipment Solutions, USA) on the beam paths 1 and 2 were used for fast mechanical scanning of the coherence time delay t_1 which was scanned from -3.06 ps to 3.06 ps in a continuous manner, at a rate which almost sample at the Nyquist frequency of the mid-IR field [234]. Due to the diverse wavevector dependence of the rephasing (RE) and non-

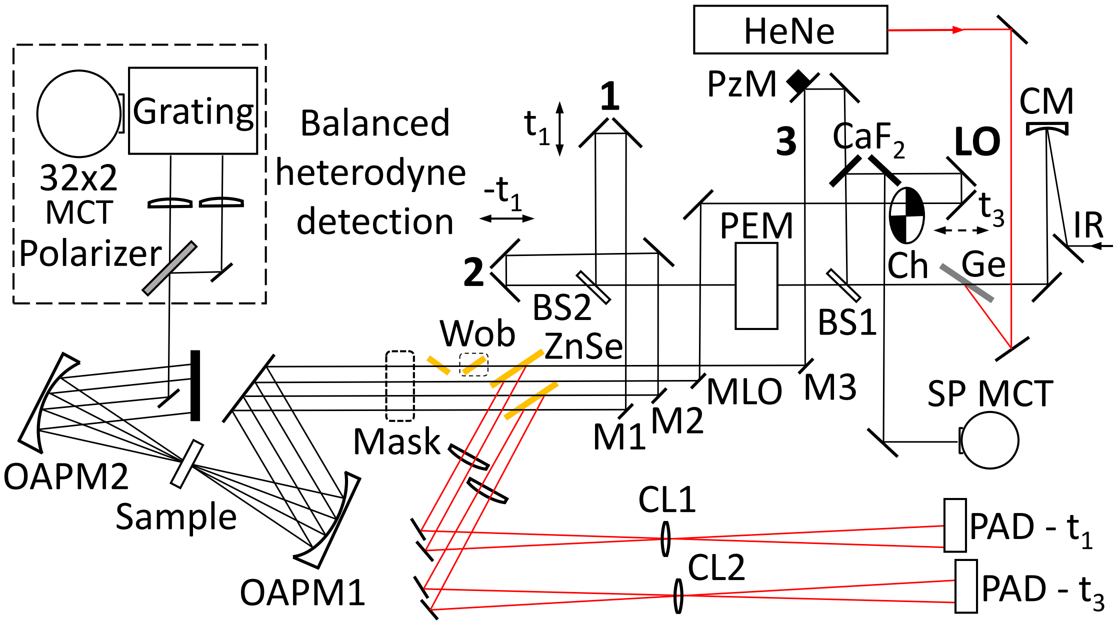


Figure A.1: 2D-IR experimental setup. HeNe, HeNe cw laser (λ = 632.8 nm, 05-LHP-151, Melles Griot, USA); CM, curved mirror (f = 315 mm); Ge, germanium plate set at the Brewster angle; Ch, optical chopper (used only to set the two phase modulators PEM and Wob for the exact-quasi-phase-cycling sequence and not while measuring the 2D-IR spectra); LO, local oscillator; BS1-BS2, custom made BaF$_2$ beam splitters (R:T, BS1 = 33:67, BS2 = 50:50); SP MCT, single pixel mercury cadmium telluride detector; PzM, piezo-mounted mirror; PEM, photo elastic modulator; ZnSe, ZnSe windows at the Brewster angle; M1-M2-M3-MLO, mirrors of the beams 1-2-3 and local oscillator; CL1-CL2, cylindrical lenses; PAD-t$_1$-PAD-t$_3$, photodiode array detectors for HeNe interference tracking of the coherence times t$_1$ and t$_3$; Wob, wobbling ZnSe window device at the Brewster angle; Mask, special mask equipped with three $\lambda/2$ waveplates, two servos and a polarizer as described in Fig. A.2 ; OAPM1-OAPM2, off-axis parabolic mirrors (f = 101.6 mm); 32x2 MCT, 2x32 pixels mercury cadmium telluride array detector.

rephasing (NR) contributions to the 2D-IR signal which are emitted in different directions in the box-CARS geometry experiment ($k_{2DIR} = \mp k_1 \pm k_2 + k_3$, where $k_{RE} = -k_1 + k_2 + k_3$ and $k_{NR} = +k_1 - k_2 + k_3$) [235], the RE and NR interferograms were acquired independently, scanning the voice coil stages in turns. The exchange of the time ordering between the first two pulses (beams 1 and 2) allowed the use of the same detection scheme for both RE ($t_1 > 0$) and NR ($t_1 < 0$) data collections. Two ZnSe windows set at the Brewster angle were employed to split the reflected s-polarized HeNe from the transmitted p-polarized mid-IR beams. After focusing the two couples of reflected beams (1-2 and 3-LO) on two cylindrical lenses, the HeNe fringes were generated. The interference patterns were then tracked using two photodiode array detectors (Hamamatsu) for different purposes. The interference between the beams 1 and 2 set the coherence time t_1, which ensured the binning of the photon echo interferograms with a time resolution of $\approx$ 2.11 fs (HeNe wave period). The one between the beams 3 and LO set instead the coherence time t_3, which was kept fixed (within half of the HeNe wave period) for phase stabilization [236]. The latter was performed by a feedback controlled piezo mirror mounted on the beam path 3, which was actively controlled using the t_3 coherence time photodiode array detector (PAD-t_3 in Fig. A.1). The combination of two synchronized phase modulators, a photoelastic modulator (Hinds Instruments, USA) on the beam path 1-2 and a wobbling Brewster ZnSe window device (Electro-Optical Products Corp., USA) on the beam path 3, ensured the complete scattering suppression *via* exact-quasi-phase-cycling (see Tab. A.1), which was implemented at the repetition rate of $\approx$ 5.13 kHz, while the original implementation of Bloem *et al.* [231] was performed at $\approx$ 1 kHz. An additional ZnSe window of the same dimensions as the one mounted on the wobbling device (15 x 10 x 3 mm) was added to the beam path 3 at minus the Brewster angle to restore the box-CARS geometry of the 2D-IR experiment (see Fig. A.1). Then, a special mask was introduced (see Fig. A.2 for a detailed representation), where the polarizations of the beams 1 and 2 were controlled with the help of two $\lambda/2$ waveplates (B. Halle, 4.75 μm) connected independently to two servos (Spektrum A5040 Digital-Servo). Thanks to the box-CARS geometry of the experiment, the intensity of the local oscillator could be adjusted apart from the ones of the other three beams, using the combination of a $\lambda/2$ waveplate (B. Halle, 4.75 μm) followed by a BaF_2 holographic wire grid polarizer (WP25H-B, Thorlabs). After the mask, the beams 1, 2, 3, and LO were focused on the sample employing an off-axis parabolic mirror (f = 101.6 mm). The emitted 2D-IR signal was dispersed on a 150 groove/mm grating blazed at 4 μm (HORIBA Scientific, France) and detected using a 2x32 pixels MCT array detector (Infrared Associates, USA).

To acquire the pump-induced linear dichroism spectra (see paragraph A.4.1), the polarizations of the beams 1 and 2 were set to be both s- or p-polarized, while the ones of the beams 3 and LO were kept fixed p- and s-polarized, respectively. For isotropic symmetry samples in these conditions, the polarization of the emitted 2D-IR signal is p-polarized [232], which is compatible with the balanced heterodyne detection scheme proposed by Hamm and Zanni [237], where a polarizer is placed at 45° respect to both the s-polarizad LO and the p-polarized 2D-IR signal.

To set the spatial and the temporal overlap among the four pulses, a 10 μm pinhole was placed at the sample position and employed as a scattering source. The temporal interferograms of the scattered beams 1-2 and 3-LO, were used to find the time zero and to minimize the phase difference $\Delta\phi_{1,2}$ - $\Delta\phi_{3,LO}$ between the two pairs of pulses as described by Backus *et al.* [238]. The 2D-IR spectra for the two different pump excitations (beams 1-2 both p- or s-polarized, parallel $\|$ and

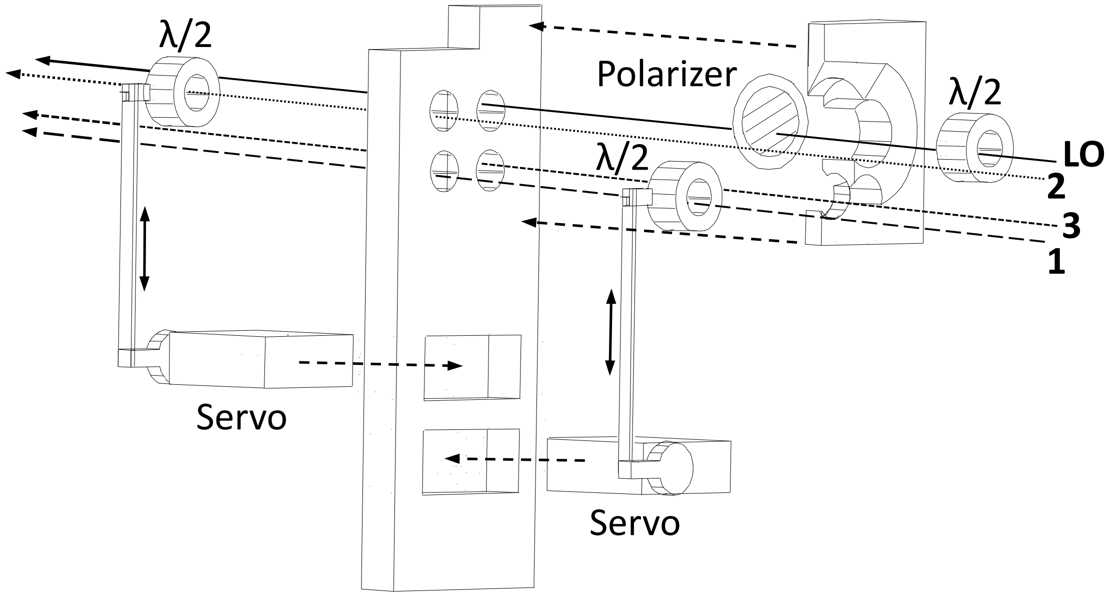


Figure A.2: Schematic drawing of the special mask with three $\lambda/2$ waveplates (B. Halle, 4.75 μm), two servos (Spektrum A5040 Digital-Servo) and a BaF_2 holographic wire grid polarizer (WP25H-B, Thorlabs) mounted on it. Each of the two servos is connected to a $\lambda/2$ waveplates, and both are controlled independently to set the beam 1 and 2 polarizations to be both s- or p-polarized for the different 2D-IR experiments, while the beam 3 is always p-polarized. The intensity of the local oscillator can be adjusted with the combination of a $\lambda/2$ waveplate, followed by a polarizer. The polarization of the local oscillator is set to be s-polarized for the balanced heterodyne detection scheme.

perpendicular $\perp$ data acquisitions, respectively) were then phased independently employing the method proposed by Johnson *et al.* [230] where the phase offset introduced by the D_2O background was extracted directly from the experimental data, compensating for it. The $||$ and $\perp$ pump-excited spectra were acquired quasi-simultaneously, switching the polarizations of the beams 1 and 2 stepwise from both p- to s-polarized in a synchronized manner with the help of the two servos. The switching time step was set by a feedback loop coming from the acquisition card (NI6225 National Instruments, USA) based on an oversampling binning condition of the acquired interferograms, which had to be fulfilled ($\approx$ 7-8 bins per mid-IR fringe at 2110 cm^{-1}). The servos were switched every *circa* 5-10 seconds using the Micro Maestro 6-Channel USB Servo Controller (Pololu, USA) which could be easily controlled by LabVIEW.

To set the exact sequence for the complete scattering suppression (see Tab. A.1), the two phase modulators (PEM for the beams 1-2 and Wobbler for the beam 3) had to be set independently in a synchronized manner over four consecutive laser shots. The only beam available which was not affected by any phase modulation was the LO and it was exploited together with a 10 μm pinhole placed at the sample

Table A.1: Exact-quasi-phase-cycling sequence as described by Bloem *et al.* [231] for complete scattering suppression. Two phase modulators, a photo elastic modulator (PEM) and a wobbling Brewster ZnSe window device (Wobbler) were used to modulate on a shot-to-shot basis the phases of the beams 1-2 and 3, respectively. As a result of it, the 2D-IR signals (E_{2DIR}) after the 1st and 3rd shot modulations have the same sign but opposite respect to the case after the 2nd and the 4th shot modulations.

	Laser shot			
	1	2	3	4
$\phi^*_{1,2}$ (PEM)	π^*	π^*	0	0
ϕ^*_3 (Wobbler)	0	π^*	0	$-\pi^*$
$\phi_{E^*_{2DIR}} = \phi^*_1 + \phi^*_2 + \phi^*_3$	$2\pi^*$	$3\pi^*$	0	$-\pi^*$
$E_{2DIR} \propto cos(\omega t + \phi_{E^*_{2DIR}})$	$+$	$-$	$+$	$-$

position as a source of scattering to promote the interference with the other scattered beams. Due to the fact that the intensity of the LO was much smaller than the ones of the other three beams ($\approx 1\%$) the help of a chopper (MC2000B optical chopper mounted with MC2F57B chopper blade, Thorlabs) on the LO beam path was needed to remove the big static background contribution of the beams 1, 2 or 3 to the interference pattern with the LO. Acting in this way, the relative phase differences among the beams 1, 2 and 3 could be set correctly. It has to be stressed that the chopper was used only to set the two phase modulators properly to achieve the exact-quasi-phase-cycling sequence [231] (see Tab. A.1) and not while measuring the 2D-IR spectra.

A.4.1 Pump-Induced Linear Dichroism Spectra

The acquisition of 2D-IR spectra without being affected by any scattering contamination is a great challenge *per se* [17, 231]. The high sensitivity of the 2D-IR spectrometer in the box-CARS geometry, ensured by the possibility to control the local oscillator (LO) and the beams 1, 2 and 3 independently of each other, combined with the option to collect $\parallel$ and $\perp$ pump-excited spectra quasi-simultaneously, has outstanding potential. These advantages can be exploited together to obtain site-specific information in biological samples, employing the IR label azidohomoalanine (Aha) [35, 36, 99–103] and suppressing to a great extent the D_2O absorption background from the Aha absorption signal as described below.

The sample employed for testing the updated version of the 2D-IR spectrometer in the box-CARS geometry was a solution of azidohomoalanine (Aha) in D_2O

(volume needed ≈ 0.5 μL, Aha concentration ≈ 1.2 mM) placed in a static cell consisting of two 2 mm thick CaF_2 windows (Crystran Ltd, UK) separated by a 25 μm Teflon spacer. The data were collected with both $\parallel$ and $\perp$ pump excitations as described before at different waiting times ($T_2 = 200$ fs, 300 fs, 400 fs and 500 fs) obtaining two indipendent datasets for each T_2 (see Fig. A.3 and Fig. A.4). The subtraction of the two spectral datasets ($\parallel - \perp$ pump-excited) gives rise to the pump-induced linear dichroism spectra, where the D_2O background absorption was suppressed to a variable extent, and the suppression increases with T_2 (see Fig. A.4, Fig. A.7, and Tab. A.2 for details). The time needed for both $\parallel$ and $\perp$ data acquisitions at each T_2 was *circa* 25 min ($\approx 7.7\text{x}10^6$ laser shots) with a laser noise below 0.3% RMSD over 500 consecutive laser shots.

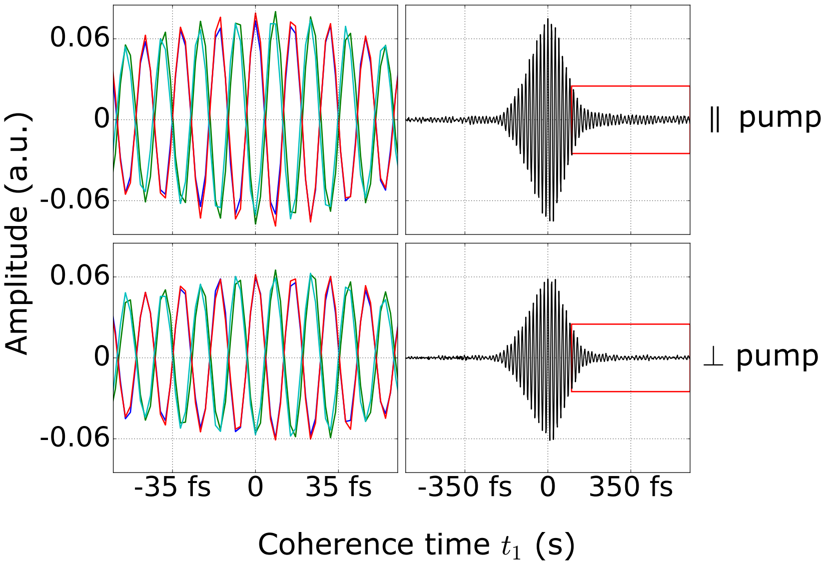


Figure A.3: Photon echo signals collected at $T_2 = 200$ fs and at the detection frequency value of *circa* 2114 cm^{-1} for both $\parallel$ and $\perp$ data acquisitions (top and bottom plots, respectively). The different colours in the plots on the left highlight the exact-quasi-phase-cycling sequence over 4 consecutive laser shots where each interferogram has been acquired with a different quasi-phase-shift value (see Tab. A.1, 1st shot blue, 2nd shot green, 3rd shot red, 4th shot turquoise). The interferograms are coupled two by two (1st-3rd and 2nd-4th shots) with a relative phase difference of π which causes the change of the signal sign between the two pairs. After an appropriate linear combination of them (1st−2nd+3rd−4th shot black, right plots) the photon echo signal is not affected by any scattering contribution. The RE part at $t_1 \geq 100$ fs has been highlighted in red for both $\parallel$ and $\perp$ pump excitations to show the different extent of the Aha photon echo signal in the two cases.

To highlight the shot-to-shot exact-quasi-phase-cycling data collection, an example of the interferograms acquired with $\parallel$ and $\perp$ pump excitations are showed in Fig. A.3 (top left and bottom left plots, respectively). Due to the cosine phase dependency of the 2D-IR signal ($E_{2DIR} \propto cos(\omega t + \phi_{E^*_{2DIR}})$), the four interferograms

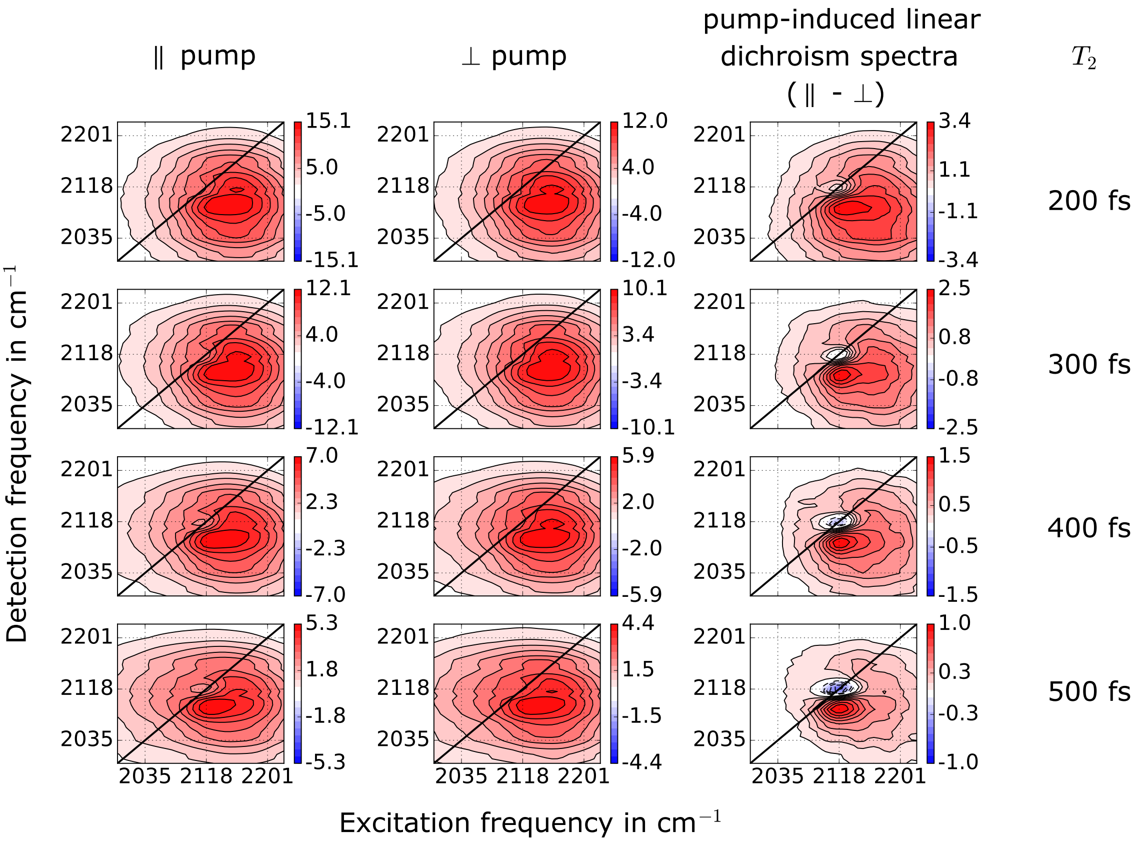


Figure A.4: 2D-IR spectra of *circa* 1.2 mM Aha in D_2O as a function of the pump polarization (|| or ⊥) and the waiting time T_2. The first and the second columns have been acquired with || and ⊥ pump excitations, respectively. The third column has been obtained by subtraction of the first two giving rise to the pump-induced linear dichroism spectra (|| − ⊥ pump-excited spectra). The subtraction suppresses partially the strong background absorption of D_2O, and the suppression increases with T_2.

collected with a different quasi-phase-shift are coupled in pairs with a relative phase difference of π (see Tab. A.1 and Fig. A.3). The appropriate linear combination of them gives rise to a photon echo interferogram for each || and ⊥ data acquisitions without any scattering contribution (see Fig. A.3, top right and bottom right plots, respectively and Tab. A.1). Each of them is the result of the sum of the D_2O and the Aha photon echo responses. The ⊥ pump-excited interferogram in Fig. A.3 has an overall lower amplitude (*circa* 20% smaller) compared to the || one, due to the lower amount of excited molecules detected by the p-polarized probe beam in the first case.

In a 2D-IR experiment, the signal strength scales quadratically with the extinction coefficient (ϵ) and linearly with both optical path length (b) and concentration (c), which is an advantage respect to FTIR spectroscopy, where it scales only linearly with ϵ, b and c ($Abs_{2DIR} = \epsilon^2 bc$, $Abs_{FTIR} = \epsilon bc$). This peculiarity enhances the Aha signal which has a relatively high extinction coefficient and low concentration ($\epsilon_{Aha} \approx$ 350-400 M^{-1} cm^{-1} [102], $c_{Aha} \approx 1.2$ mM) in comparison to D_2O which has a small extinction coefficient at *circa* 2110 cm^{-1} and a huge concentration ($\epsilon_{D_2O} \approx 1.8$ M^{-1} cm^{-1} [31, 239], $c_{D_2O} \approx 55$ M). Aha and D_2O are two oscillators with different vibrational lifetimes ($\tau_{D_2O} \approx 400 \pm 30$ fs [240], $\tau_{Aha} \approx 1000 \pm 100$ fs [102]) and anisotropic behaviours. Bulk liquid D_2O has the exceptional characteristic to have an ultrafast decay of the anisotropy which is almost complete after 200 fs (anisotropy value *circa* 0.055) due to the Förster resonance energy transfer mechanism operating among the neighboring D_2O molecules, as described by Piatkowski *et al.* [240]. On the other hand, Aha is expected to have a slower anisotropic decay due to only its rotational motion and not to an exceptional ultrafast depolarizing mechanism as in the case of D_2O. As a result of these different characteristics, a discrimination between the two oscillators has been experimentally observed. The Aha photon echo signal is clearly visible on top of the D_2O response in the RE part at $t_1 \geq 100$ fs due to its longer lifetime. It is evident for the $\|$ data acquisition, while is not anymore for the $\perp$ one (see Fig. A.3, red square), which highlights the different anisotropic behaviours between Aha and bulk liquid D_2O, already at $T_2 = 200$ fs. Analyzing the data in the frequency domain after a Fourier transformation along the pump excitation axis, the same different anisotropic behaviours have been observed for the two oscillators. Several 2D-IR spectra were acquired with T_2 delays from 200 to 500 fs (see Fig. A.4). It can be clearly observed that in all the $\|$ pump-excited spectra there is a more pronounced Aha bleaching feature on the top of the D_2O absorption around 2118 cm^{-1} along the diagonal, respect to the corresponding $\perp$ pump-excited spectra (first two columns in Fig. A.4). The relative bleaching feature of Aha increases with T_2 for both $\|$ and $\perp$ pump-excited datasets due to the longer lifetime of Aha respect to D_2O. Another minor effect has been observed increasing T_2; the D_2O background absorption shifts towards lower frequencies along the pump excitation axis, as a possible consequence of heat dissipation effects on the D_2O absorption band. The third column in Fig. A.4 represents the difference between the first two giving rise to the pump-induced linear dichroism spectra ($\| - \perp$ pump-excited spectra). It has to be stressed that this subtraction decreases quantitatively the D_2O absorption background in favor of the Aha signal, enhancing its detectivity. This enhancement increases with T_2 due to the different lifetimes of the two oscillators and to the *de facto* complete anisotropic decay of bulk liquid D_2O at $T_2 = 500$ fs (anisotropy value *circa* 0.01 [240]). To reveal this enhancement effect more clearly, the series of 2D-IR spectra acquired at $T_2 = 200$ and 500 fs with $\|$ and $\perp$ pump excitations are shown in Fig. A.5 and Fig. A.6, while the corresponding pump-induced linear dichroism spectra in Fig. A.7. A section of the 2D-IR spectra along the detection frequency axis at $\tilde{\nu}_3 \approx 2121$ cm^{-1} has been highlighted, where the best visual discrimination of the Aha and the D_2O signals is present, due to the different signs of the Aha absorption bleaching (negative) and the D_2O excited state absorption (positive). To quantify the extent of the D_2O background suppression in each pump-induced linear dichroism spectrum, an estimate of the Aha contribution to the total signal (Aha + D_2O background) has been made, as shown and described in Fig. A.7, and Tab. A.2.

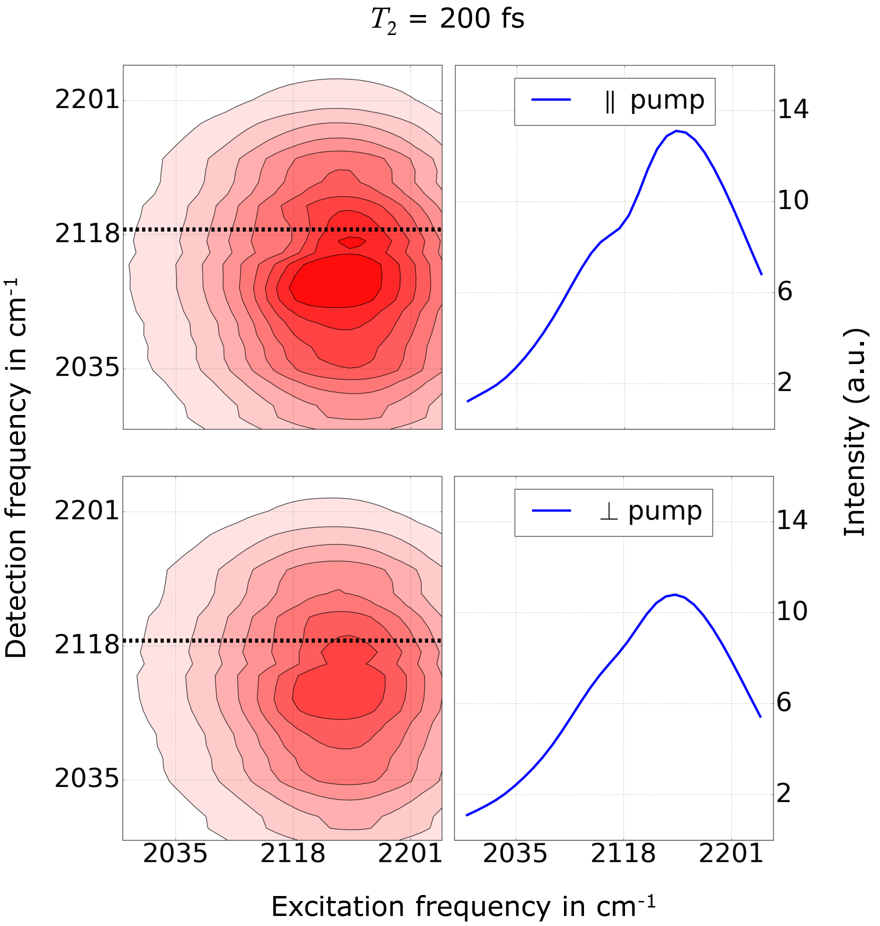


Figure A.5: 2D-IR spectra of *circa* 1.2 mM Aha in D_2O acquired with $\|$ and $\perp$ pump excitations at $T_2 = 200$ fs (top left and bottom left graphs, respectively). The spectra on the right side represent the sections along the detection frequency axis at $\tilde{\nu}_3 \approx 2121$ cm^{-1} which have been highlighted with the black dotted lines on the corresponding 2D-IR spectra. The color scale of the 2D-IR spectra are relative to the spectrum collected with the $\|$ pump excitation (top left graph).

In conclusion, the 2D-IR spectrometer in the box-CARS geometry was described in details with its latest improvements: the exact-quasi-phase-cycling scattering suppression [231] implemented at the repetition rate of $\approx$ 5.13 kHz, and the polarization control of the pump excitation by a home-built special mask where two servos were connected to two $\lambda/2$ waveplates (see Fig. A.2). The combination of these new characteristics of the optical setup together with the fast mechanical scanning of

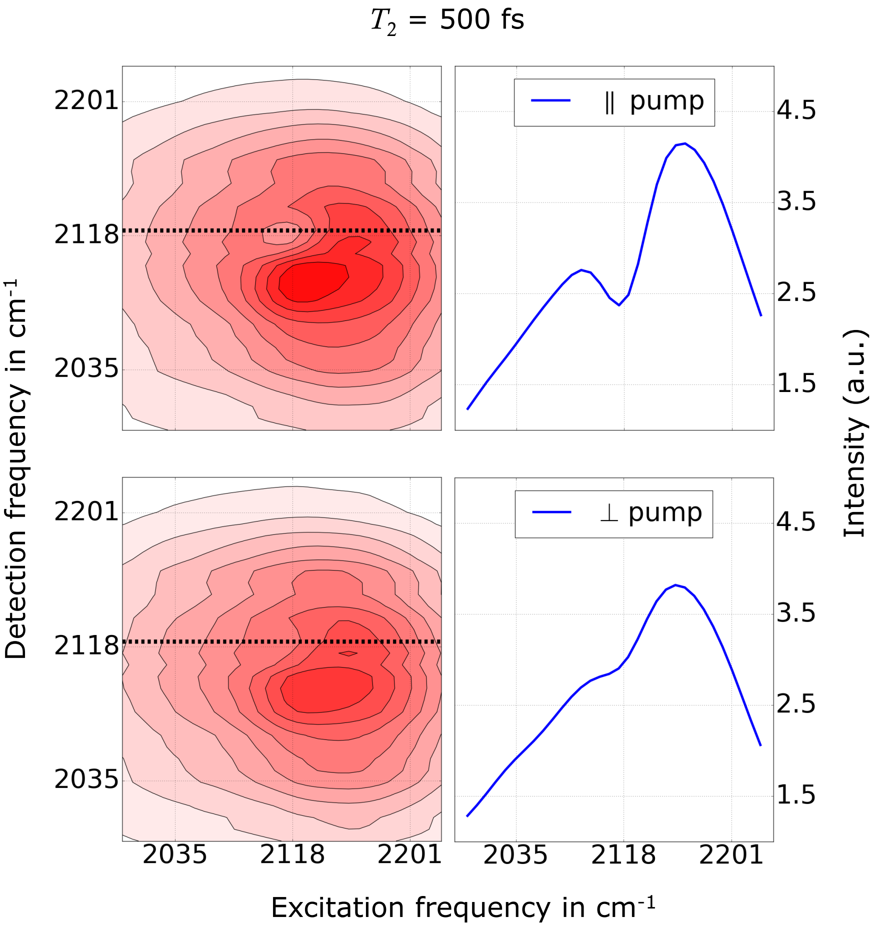


Figure A.6: 2D-IR spectra of *circa* 1.2 mM Aha in D_2O acquired with $\parallel$ and $\perp$ pump excitations at $T_2 = 500$ fs (top left and bottom left graphs, respectively). The spectra on the right side represent the sections along the detection frequency axis at $\tilde{\nu}_3 \approx 2121$ cm^{-1} which have been highlighted with the black dotted lines on the corresponding 2D-IR spectra. The color scale of the 2D-IR spectra are relative to the spectrum collected with the $\parallel$ pump excitation (top left graph).

the coherence time delay t_1 permitted the excellent quality, quasi-simultaneous data collection of the $\parallel$ and the $\perp$ pump-excited spectra. The subtraction of the two datasets ($\parallel - \perp$ pump-excited at each T_2) gave rise to the pump-induced linear dichroism spectra (see the spectra in Fig. A.4 – 3rd column, and in Fig. A.7). As a result of the different anisotropic behaviours and lifetimes of the two oscillators, Aha and bulk liquid D_2O, a substantial reduction of the D_2O background has been observed increasing the waiting time T_2 from 200 to 500 fs. The contribution of the Aha signal relative to the total signal (Aha + D_2O background) increased from $\approx 38\%$ to $\approx 76\%$ (see Tab. A.2 for details) and its enhancement is clearly visible, as shown in the spectra of Fig. A.4 – 3rd column, and of Fig. A.7. Moreover, from the comparison of the pump-induced linear dichroism spectrum of Aha acquired at

Table A.2: Estimate of the contribution of the Aha signal in % relative to the total signal (Aha + D_2O background) in the pump-induced linear dichroism spectra (see Fig. A.4 and Fig. A.7) as a function of the waiting time T_2. The values have been calculated by a relative quantification of the D_2O and Aha signals at $\tilde{\nu}_3 \approx 2121$ cm^{-1} as shown in Fig. A.7, where the two signals have positive and negative signs, respectively. A linear fit of the experimental data has been performed without considering the Aha bleaching signal to make the estimate mentioned above. The error associated with each estimate is smaller than 1%.

T_2 (fs)	Pump Polarization	Aha signal relative to the Total signal (Aha + D_2O background) in %
200	$\parallel$ - $\perp$	38 ± 0.4
300	$\parallel$ - $\perp$	51 ± 0.5
400	$\parallel$ - $\perp$	63 ± 0.6
500	$\parallel$ - $\perp$	76 ± 0.8

$T_2 = 500$ fs with the D_2O background-subtracted Aha spectrum in Fig. A.8, we can observe that both spectra have similar features with the same Aha bleaching absorption frequency at ≈ 2118 cm^{-1}. Therefore, we can conclude that the collection of pump-induced linear dichroism spectra of Aha sample solutions at $T_2 = 500$ fs is a useful tool to obtain information about the polarity changes of the Aha chemical environment, which can be investigated from a spectroscopic point of view, observing the frequency shifts of the Aha bleaching absorption. Due to the extensive employment of Aha in Biophysics as a sensitive, site-specific, and minimally invasive IR probe [35, 36, 101–103], it is crucial to be able to perform spectroscopic experiments on Aha using the least amount of sample as possible, especially when the objects of study are sophisticated proteins which synthesis, purification and sample preparation are a significant effort. Hence, with the presented 2D-IR spectrometer in the box-CARS geometry a static cell with only ≈ 0.5 μL of ≈ 1.2 mM of Aha sample solution is needed to perform the 2D-IR experiments with the remarkable advantage to using small amounts of sample in the sub-nmol regime.

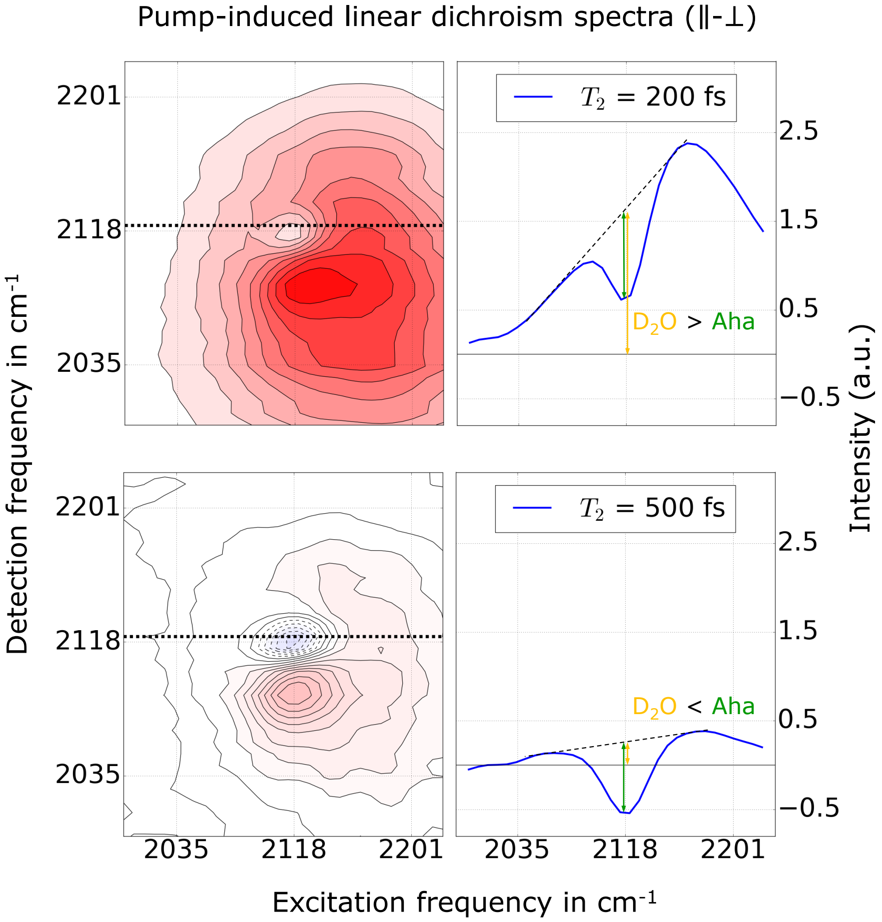


Figure A.7: Pump-induced linear dichroism spectra ($\| - \perp$ pump-excited spectra) of *circa* 1.2 mM Aha in D_2O acquired at $T_2 = 200$ and 500 fs (top left and bottom left graphs, respectively). The spectra on the right side represent the sections along the detection frequency axis at $\tilde{\nu}_3 \approx 2121$ cm^{-1} which have been highlighted with the black dotted lines on the corresponding 2D-IR spectra. Each section has been linearly fitted (black dashed line) excluding the Aha bleaching signal from the fitting procedure in order to quantify the relative contributions of the Aha (negative) and D_2O (positive) signals to the total signal (see Tab. A.2 for details). The green and orange arrows in the graphs represent the absolute values of the Aha and the D_2O contributions, respectively. The color scale of the 2D-IR spectra are relative to the spectrum collected at $T_2 = 200$ fs (top left graph).

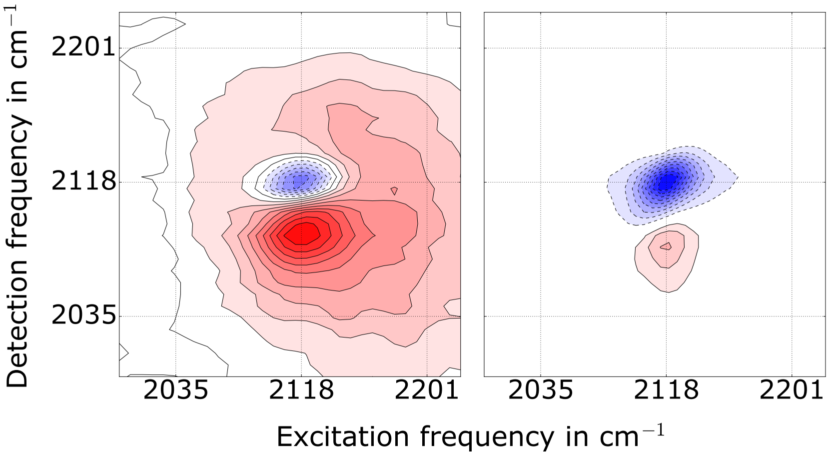


Figure A.8: Pump-induced linear dichroism spectrum ($\parallel - \perp$ pump-excited spectra) of *circa* 1.2 mM Aha in D_2O acquired at $T_2 = 500$ fs (left graph). The right graph represents the spectrum of *circa* 2.5 mM Aha in D_2O acquired at $T_2 = 300$ fs with $\parallel$ pump excitation, where the Aha signal has been isolated after subtracting the D_2O background signal, which has been acquired separately in an analogous manner. The 2D-IR spectra have the same number of contour lines but the color scales are different in the two cases.

List of Figures

List of Tables

Abbreviations

E_a activation energy. 89, 93

K_D dissociation constant. 57, 59–61, 68, 90, 91, 93–96, 112, 114

Φ hydrophobic amino acid residue, usually valine (V), isoleucine (I) or leucine (L). 11, 12

$\Phi_{\boldsymbol{cis \rightarrow trans}}$ *cis* to *trans* isomerization quantum yield. 15

$\Phi_{\boldsymbol{trans \rightarrow cis}}$ *trans* to *cis* isomerization quantum yield. 15

ϵ extinction coefficient or absorption cross section. 13, 15, 105

λ_{max} maximum absorption wavelength. 15

b optical path length. 105

c concentration. 105

k_B Boltzmann constant. 89

k_{off} kinetic dissociation constant. 62, 91, 92

k_{on} kinetic association constant. 91

$t_{1/2}$ half-life. 91, 96

—$\mathbf{SO_3}^-$ sulfonate group. 17, 18, 41, 59

Å Ångström. 9, 14, 58

2D-IR two-dimensional infrared. 5–7, 21, 41, 57, 89, 99–111, 119

A alanine. 59

Aha azidohomoalanine. 5, 9, 13, 21, 41, 57, 89, 99, 103, 105, 106, 108–111

C cysteine. 16, 18–20, 59

CD circular dichroism. 16, 60, 61, 94, 112, 114

cw continuous wave. 21, 94, 96, 99

D aspartic acid. 19

DLG1 *Drosophila* disc large tumor suppressor. 9

E glutamic acid. 12, 59

F phenylalanine. 12, 19

Fmoc fluorenylmethyloxycarbonyl. 13

FTIR Fourier Transform Infrared. 86, 87

FWHM Full Width at Half Maximum. 99

G glycine. 9

H histidine. 19, 20

I isoleucine. 11

IR infrared. 5–7, 13, 85, 89, 95, 96, 103, 109

ITC isothermal titration calorimetry. 60, 61, 93, 94, 114

K lysine. 9, 12, 19, 20, 59

K38Aha Aha-mutated Lysine-38 PDZ2 domain. 41, 90

L leucine. 9, 11, 12

LO local oscillator. 99, 103

M methionine. 12, 13

MAGUKs membrane-associated guanylate kinases. 9, 92

MCT mercury cadmium telluride detector. 95, 99

MD molecular dynamics. 41, 67, 91

MEM maximum entropy method. 97, 98

NMR Nuclear Magnetic Resonance. 6, 8, 58, 90

NR non-rephasing. 99, 100

OPA optical parametric amplifier. 95, 99

P proline. 19, 20

PDB Protein Data Bank. 20

PEM photo elastic modulator. 103

PSD-95 postsynaptic density protein 95. 9, 92

Q glutamine. 19, 59

R arginine. 9, 19, 59

RE rephasing. 99, 100

RMSD root-mean-square deviation. 9, 104

RNA ribonucleic acid. 18, 19

RNase A ribonuclease A. 9, 18–20

RNase S ribonuclease S. 5, 8, 9, 17, 19, 20, 68, 85, 87, 88, 91, 92, 119

S serine. 11, 12, 59

S-pep(6,13) S-peptide photoswitchable variant with spacing $i+7$; the photoswitch anchoring amino acids are in positions 6 and 13 of the S-peptide primary structure. 68, 85–88, 91, 95, 96

T threonine. 11, 12

TRIR transient infrared. 5, 6, 8, 57, 58, 89, 90, 93

V valine. 11, 12, 59

VIS visible. 6, 14, 16, 17, 89, 95

W tryptophan. 59–61, 93, 94, 112, 114

X any amino acid. 9, 11

Y tyrosine. 92, 93

ZO-1 zonula occludens-1. 9

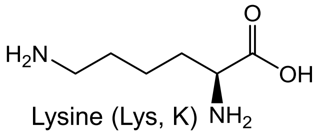
Acknowledgements

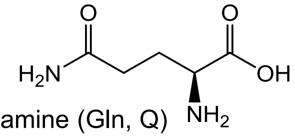
When I moved to Zurich in February 2015, the first two things that impressed me were the cold winter and the perennial snow on one side and the amazing University on the other. Despite the difficulties encountered in adapting to the new environment, what always motivated me during my PhD studies was the possibility to conduct scientific research at the highest level, thanks to the combination of the high scientific knowledge of the people working in the Hamm group together with the incredible resources provided by the University of Zurich and Switzerland.

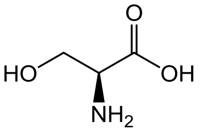
I would like to thank the following people:

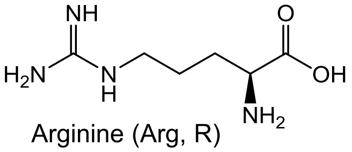
- Prof. Dr. Peter Hamm for the possibility and trust given me to work on the challenging "protein project", interesting discussions and suggestions.
- Dr. Brankica Janković and Olga Božović for the huge amount of time spent together to solve protein related problems and for the funny moments spent together in our office/balcony. Without them, the research activity performed and explained in this thesis could not have been possible.
- Ricardo J. Fernández-Terán for his constant availability to solve laser lab and programming issues, for scientific discussions and for talking Spanish together.
- Dr. Jan Helbing for interesting discussions, explanations, suggestions, and help in the laser labs.
- Gökçen Tek for his Mediterranean attitude typical of a Turkish guy.
- Arian Berger for his friendliness.
- Gustavo Ciardi for reminding me how to speak "Fiorentino" (the language of Florence).
- Roland Zehnder for his constant help to provide me home-made mechanical parts whenever I needed for daily routine work in the laser labs. I defined him as "the magician of metal and plastic".

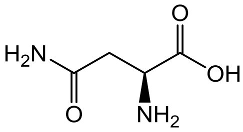
- Rolf Pfister for his help in the synthesis and purification of peptides and photoswitches.
- Prof. Dr. Gerhard Stock and coworkers for their help regarding the simulation part of the PDZ2 domain and RNase S papers.
- Prof. Dr. Oliver Zerbe to be part of my PhD committee and for useful discussions/suggestions.
- Prof. Dr. Ben Schuler for the possibility to use the equipment needed for protein expression and for useful discussions/suggestions.
- Dr. Serge Chesnov and the Functional Genomics Center Zurich (FGCZ) for mass and amino acids analysis.
- Dr. Klemens L. Koziol and Dr. Philip J. M. Johnson for their introduction to the 2D-IR spectrometer in the box-CARS geometry of lab 3.
- Dr. Brigitte Stucki-Buchli for her introduction to the “protein project” on the PDZ2 domain, which has been taken over by me, Olga Božović, and Dr. Brankica Janković.
- Dr. Lukas Frey for his introduction to protein expression during the first month of my PhD.
- The other group members of the Hamm group for the overall time spent together during the PhD.
- The Graduate School of Chemical and Molecular Sciences Zurich (CMSZH) for monthly aperos and yearly retreats.
- The Swiss National Science Foundation (SNF), the NCCR MUST (Molecular Ultrafast Science and Technology), and the European Research Council (ERC) Advanced Investigator Grant (DYNALLO) for funding.
- Last but not least, my parents, brother, and sister for their constant moral support and motivation given to me during these years spent in Zurich, to whom I dedicate this thesis.

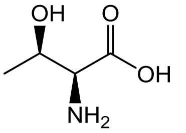

Lysine (Lys, K)

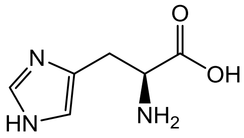

Glutamine (Gln, Q)

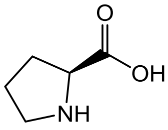

Serine (Ser, S)

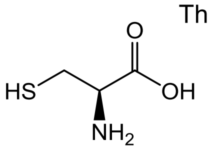

Arginine (Arg, R)

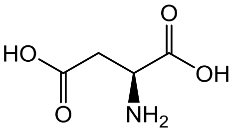

Asparagine (Asn, N)

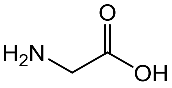

Threonine (Thr, T)

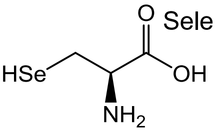

Histidine (His, H)

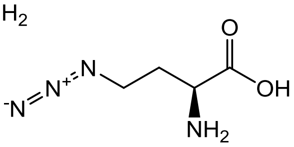

Proline (Pro, P)

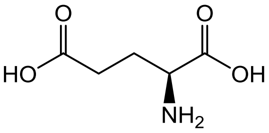

Cysteine (Cys, C)

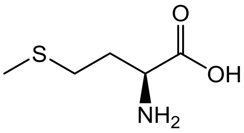

Aspartic acid (Asp, D)

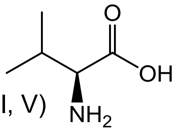

Glycine (Gly, G)

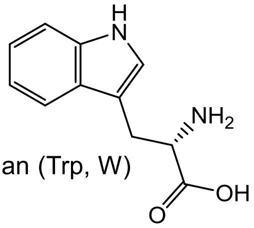

Selenocysteine (Sec, U)

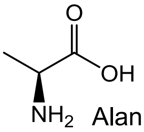

Azidohomoalanine (Aha)

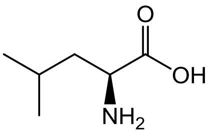

Glutamic acid (Glu, E)

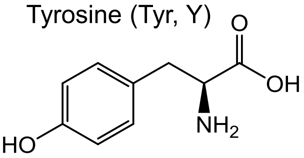

Methionine (Met, M)

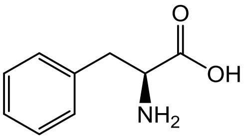

Valine (Val, V)

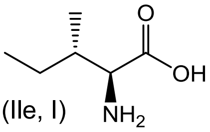

Tryptophan (Trp, W)


Alanine (Ala, A)


Leucine (Leu, L)


Tyrosine (Tyr, Y)


Phenylalanine (Phe, F)


Isoleucine (Ile, I)

Bibliography

1. Noda, I. Two-dimensional infrared spectroscopy. *Journal of the American Chemical Society* **111,** 8116–8118 (1989).
2. Hamm, P., Lim, M. & Hochstrasser, R. M. Structure of the Amide I Band of Peptides Measured by Femtosecond Nonlinear-Infrared Spectroscopy. *The Journal of Physical Chemistry B* **102,** 6123–6138 (1998).
3. Ghosh, A., Ostrander, J. S. & Zanni, M. T. Watching Proteins Wiggle: Mapping Structures with Two-Dimensional Infrared Spectroscopy. *Chemical Reviews* **117,** 10726–10759 (2017).
4. Hamm, P. & Zanni, M. in *Concepts and Methods of 2D Infrared Spectroscopy* (Cambridge University Press, 2011).
5. Hochstrasser, R. M. Two-dimensional spectroscopy at infrared and optical frequencies. *Proceedings of the National Academy of Sciences* **104,** 14190–14196 (2007).
6. Bredenbeck, J., Helbing, J., Kolano, C. & Hamm, P. Ultrafast 2D IR Spectroscopy of Transient Species. *ChemPhysChem* **8,** 1747–1756 (2007).
7. Woutersen, S. & Hamm, P. Structure Determination of Trialanine in Water Using Polarization Sensitive Two-Dimensional Vibrational Spectroscopy. *The Journal of Physical Chemistry B* **104,** 11316–11320 (2000).
8. Helbing, J. & Hamm, P. Compact implementation of Fourier transform two-dimensional IR spectroscopy without phase ambiguity. *J. Opt. Soc. Am. B* **28,** 171–178 (Jan. 2011).
9. Thämer, M., De Marco, L., Ramasesha, K., Mandal, A. & Tokmakoff, A. Ultrafast 2D IR spectroscopy of the excess proton in liquid water. *Science* **350,** 78–82 (2015).
10. Fayer, M. Dynamics of Liquids, Molecules, and Proteins Measured with Ultrafast 2D IR Vibrational Echo Chemical Exchange Spectroscopy. *Annual Review of Physical Chemistry* **60,** 21–38 (2009).

11. Perakis, F. & Hamm, P. Two-Dimensional Infrared Spectroscopy of Supercooled Water. *The Journal of Physical Chemistry B* **115,** 5289–5293 (2011).
12. Kim, Y. S. & Hochstrasser, R. M. Chemical exchange 2D IR of hydrogen-bond making and breaking. *Proceedings of the National Academy of Sciences* **102,** 11185–11190 (2005).
13. Bredenbeck, J., Helbing, J. & Hamm, P. Labeling Vibrations by Light: Ultrafast Transient 2D-IR Spectroscopy Tracks Vibrational Modes during Photoinduced Charge Transfer. *Journal of the American Chemical Society* **126,** 990–991 (2004).
14. Cahoon, J. F., Sawyer, K. R., Schlegel, J. P. & Harris, C. B. Determining Transition-State Geometries in Liquids Using 2D-IR. *Science* **319,** 1820–1823 (2008).
15. Bian, H. *et al.* Ion clustering in aqueous solutions probed with vibrational energy transfer. *Proceedings of the National Academy of Sciences* **108,** 4737–4742 (2011).
16. Kraack, J. P., Frei, A., Alberto, R. & Hamm, P. Ultrafast Vibrational Energy Transfer in Catalytic Monolayers at Solid-Liquid Interfaces. *The Journal of Physical Chemistry Letters* **8,** 2489–2495 (2017).
17. Shim, S.-H., Strasfeld, D. B., Ling, Y. L. & Zanni, M. T. Automated 2D IR spectroscopy using a mid-IR pulse shaper and application of this technology to the human islet amyloid polypeptide. *Proceedings of the National Academy of Sciences* **104,** 14197–14202 (2007).
18. Kim, Y. S. & Hochstrasser, R. M. Applications of 2D IR Spectroscopy to Peptides, Proteins, and Hydrogen-Bond Dynamics. *The Journal of Physical Chemistry B* **113,** 8231–8251 (2009).
19. Rubtsov, I. V., Wang, J. & Hochstrasser, R. M. Dual-frequency 2D-IR spectroscopy heterodyned photon echo of the peptide bond. *Proceedings of the National Academy of Sciences* **100,** 5601–5606 (2003).
20. Asplund, M. C., Zanni, M. T. & Hochstrasser, R. M. Two-dimensional infrared spectroscopy of peptides by phase-controlled femtosecond vibrational photon echoes. *Proceedings of the National Academy of Sciences* **97,** 8219–8224 (2000).
21. Middleton, C. T. *et al.* Two-dimensional infrared spectroscopy reveals the complex behaviour of an amyloid fibril inhibitor. *Nature Chemistry* **4,** 355–360 (2012).

22. Kuroda, D. G. *et al.* Snapshot of the equilibrium dynamics of a drug bound to HIV-1 reverse transcriptase. *Nature Chemistry* **5,** 174–181 (Mar. 2013).
23. Herbst, J., Heyne, K. & Diller, R. Femtosecond Infrared Spectroscopy of Bacteriorhodopsin Chromophore Isomerization. *Science* **297,** 822–825 (2002).
24. Patricia, T. T. *et al.* Transient infrared spectroscopy: a new approach to investigate valence tautomerism. *Phys. Chem. Chem. Phys.* **14,** 1038–1047 (2 2012).
25. Hamm, P., Ohline, S. M. & Zinth, W. Vibrational cooling after ultrafast photoisomerization of azobenzene measured by femtosecond infrared spectroscopy. *The Journal of Chemical Physics* **106,** 519–529 (1997).
26. Anfinrud, P. A., Han, C. & Hochstrasser, R. M. Direct observations of ligand dynamics in hemoglobin by subpicosecond infrared spectroscopy. *Proceedings of the National Academy of Sciences* **86,** 8387–8391 (1989).
27. Wüthrich, Kurt. NMR with Proteins and Nucleic Acids. *Europhys. News* **17,** 11–13 (1986).
28. Sattler, M., Schleucher, J. & Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Progress in Nuclear Magnetic Resonance Spectroscopy* **34,** 93–158 (1999).
29. Dyson, H. J. & Wright, P. E. Unfolded Proteins and Protein Folding Studied by NMR. *Chemical Reviews* **104,** 3607–3622 (2004).
30. Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **274,** 1531–1534 (1996).
31. Koziol, K. L., Johnson, P. J., Stucki-Buchli, B., Waldauer, S. A. & Hamm, P. Fast infrared spectroscopy of protein dynamics: advancing sensitivity and selectivity. *Current Opinion in Structural Biology* **34,** 1–6 (2015).
32. Akke, M. NMR methods for characterizing microsecond to millisecond dynamics in recognition and catalysis. *Current Opinion in Structural Biology* **12,** 642–647 (2002).
33. Kubo, R. in *Advances in Chemical Physics* 101–127 (1969).
34. Hamm, P. & Zanni, M. in *Concepts and Methods of 2D Infrared Spectroscopy* 145–165 (Cambridge University Press, 2011).
35. Zanobini, C. *et al.* Azidohomoalanine: A Minimally Invasive, Versatile, and Sensitive Infrared Label in Proteins To Study Ligand Binding. *The Journal of Physical Chemistry B* **122,** 10118–10125 (2018).

36. Taskent-Sezgin, H. *et al.* Azidohomoalanine: A conformationally sensitive IR probe of protein folding, protein structure, and electrostatics. *Angew. Chem. Int. Ed.* **49,** 7473–7475 (2010).

37. Thielges, M. C. *et al.* Two-dimensional IR spectroscopy of protein dynamics using two vibrational labels: a site-specific genetically encoded unnatural amino acid and an active site ligand. *J. Phys. Chem. B* **115,** 11294–11304 (2011).

38. Oh, K.-I., Lee, J.-H., Joo, C., Han, H. & Cho, M. β-Azidoalanine as an IR Probe: Application to Amyloid A(16-22) Aggregation. *The Journal of Physical Chemistry B* **112,** 10352–10357 (2008).

39. Getahun, Z. *et al.* Using Nitrile-Derivatized Amino Acids as Infrared Probes of Local Environment. *Journal of the American Chemical Society* **125,** 405–411 (2003).

40. Horness, R. E., Basom, E. J. & Thielges, M. C. Site-selective characterization of Src homology 3 domain molecular recognition with cyanophenylalanine infrared probes. *Anal. Methods* **7,** 7234–7241 (17 2015).

41. Van Wilderen, L. J. G. W., Kern-Michler, D., Müller-Werkmeister, H. M. & Bredenbeck, J. Vibrational dynamics and solvatochromism of the label SCN in various solvents and hemoglobin by time dependent IR and 2D-IR spectroscopy. *Phys. Chem. Chem. Phys.* **16,** 19643–19653 (2014).

42. Koziński, M., Garrett-Roe, S. & Hamm, P. 2D-IR Spectroscopy of the Sulfhydryl Band of Cysteines in the Hydrophobic Core of Proteins. *The Journal of Physical Chemistry B* **112,** 7645–7650 (2008).

43. Kim, H. & Cho, M. Infrared probes for studying the structure and dynamics of biomolecules. *Chem. Rev.* **113,** 5817–5847 (2013).

44. Waegele, M. M., Culik, R. M. & Gai, F. Site-Specific Spectroscopic Reporters of the Local Electric Field, Hydration, Structure, and Dynamics of Biomolecules. *The Journal of Physical Chemistry Letters* **2,** 2598–2609 (2011).

45. Barth, A. & Zscherp, C. What vibrations tell about proteins. *Quarterly Reviews of Biophysics* **35,** 369–430 (2002).

46. Torres, J., Briggs, J. A. & Arkin, I. T. Multiple site-specific infrared dichroism of CD3-ζ, a transmembrane helix bundle. *Journal of Molecular Biology* **316,** 365–374 (2002).

47. Kim, Y. S., Liu, L., Axelsen, P. H. & Hochstrasser, R. M. Two-dimensional infrared spectra of isotopically diluted amyloid fibrils from Aβ40. *Proceedings of the National Academy of Sciences* **105,** 7720–7725 (2008).

48. Petty, S. A. & Decatur, S. M. Intersheet rearrangement of polypeptides during nucleation of β-sheet aggregates. *Proceedings of the National Academy of Sciences* **102,** 14272–14277 (2005).
49. Shim, S.-H. *et al.* Two-dimensional IR spectroscopy and isotope labeling defines the pathway of amyloid formation with residue-specific resolution. *Proceedings of the National Academy of Sciences* **106,** 6614–6619 (2009).
50. Horness, R. E., Basom, E. J., Mayer, J. P. & Thielges, M. C. Resolution of Site-Specific Conformational Heterogeneity in Proline-Rich Molecular Recognition by Src Homology 3 Domains. *Journal of the American Chemical Society* **138,** 1130–1133 (2016).
51. Backus, E. H. G. *et al.* 2D-IR Study of a Photoswitchable Isotope-Labeled -Helix. *The Journal of Physical Chemistry B* **114,** 3735–3740 (2010).
52. Fang, C. *et al.* Two-Dimensional Infrared Spectroscopy of Isotopomers of an Alanine Rich -Helix. *The Journal of Physical Chemistry B* **108,** 10415–10427 (2004).
53. Xiong, L., Schumaker, K. S. & Zhu, J.-K. Cell Signaling during Cold, Drought, and Salt Stress. *The Plant Cell* **14,** S165–S183 (2002).
54. Pawson, T. & Nash, P. Protein-protein interactions define specificity in signal transduction. *Genes & Development* **14,** 1027–1047 (2000).
55. Bu, Z. & Callaway, D. J. Chapter 5 - Proteins MOVE! Protein dynamics and long-range allostery in cell signaling. *Advances in Protein Chemistry and Structural Biology* **83,** 163–221 (2011).
56. Nussinov, R., Tsai, C.-J. & Ma, B. The Underappreciated Role of Allostery in the Cellular Network. *Annual Review of Biophysics* **42,** 169–189 (2013).
57. Harris, B. Z. & Lim, W. A. Mechanism and role of PDZ domains in signaling complex assembly. *Journal of Cell Science* **114,** 3219–3231 (2001).
58. Lee, H. J. & Zheng, J. J. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun Signal.* **8,** 8 (2010).
59. Fuentes, E. J., Der, C. J. & Lee, A. L. Ligand-dependent dynamics and intramolecular signaling in a PDZ domain. *J. Mol. Biol.* **335,** 1105–1115 (2004).
60. Liu, X., Shepherd, T. R., Murray, A. M., Xu, Z. & Fuentes, E. J. The Structure of the Tiam1 PDZ Domain/ Phospho-Syndecan1 Complex Reveals a Ligand Conformation that Modulates Protein Dynamics. *Structure* **21,** 342–354 (2013).

61. Dhulesia, A., Gsponer, J. & Vendruscolo, M. Mapping of Two Networks of Residues That Exhibit Structural and Dynamical Changes upon Binding in a PDZ Domain Protein. *Journal of the American Chemical Society* **130,** 8931–8939 (2008).

62. Gianni, S. *et al.* Demonstration of long-range interactions in a PDZ domain by NMR, kinetics, and protein engineering. *Structure* **14,** 1801–1809 (2006).

63. Rios, P. D. L. *et al.* Functional Dynamics of PDZ Binding Domains: A Normal-Mode Analysis. *Biophysical Journal* **89,** 14–21 (2005).

64. Fuentes, E. J., Gilmore, S. A., Mauldin, R. V. & Lee, A. L. Evaluation of energetic and dynamic coupling networks in a PDZ domain protein. *J. Mol. Biol.* **364,** 337–351 (2006).

65. Kumawat, A. & Chakrabarty, S. Hidden electrostatic basis of dynamic allostery in a PDZ domain. *Proceedings of the National Academy of Sciences* **114,** E5825–E5834 (2017).

66. Palmer, A. G. NMR Characterization of the Dynamics of Biomacromolecules. *Chemical Reviews* **104,** 3623–3640 (2004).

67. Jarymowycz, V. A. & Stone, M. J. Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences. *Chemical Reviews* **106,** 1624–1671 (2006).

68. Beharry, A. A. & Woolley, G. A. Azobenzene photoswitches for biomolecules. *Chem. Soc. Rev.* **40,** 4422–4437 (8 2011).

69. Yager, K. G. & Barrett, C. J. Novel photo-switching using azobenzene functional materials. *Journal of Photochemistry and Photobiology A: Chemistry* **182,** 250–261 (2006).

70. Nägele, T., Hoche, R., Zinth, W. & Wachtveitl, J. Femtosecond photoisomerization of cis-azobenzene. *Chem. Phys. Lett.* **272,** 489–495 (1997).

71. Spörlein, S. *et al.* Ultrafast spectroscopy reveals subnanosecond peptide conformational dynamics and validates molecular dynamics simulation. *Proceedings of the National Academy of Sciences* **99,** 7998–8002 (2002).

72. Ihalainen, J. A. *et al.* Folding and unfolding of a photoswitchable peptide from picoseconds to microseconds. *Proceedings of the National Academy of Sciences* **104,** 5383–5388 (2007).

73. Szymański, W., Beierle, J. M., Kistemaker, H. A. V., Velema, W. A. & Feringa, B. L. Reversible Photocontrol of Biological Systems by the Incorporation of Molecular Photoswitches. *Chemical Reviews* **113,** 6114–6178 (2013).

74. Wyckoff, H. W. *et al.* The Structure of Ribonuclease-S at 3.5 A Resolution. *Journal of Biological Chemistry* **242,** 3984–3988 (1967).

75. Raines, R. T. Ribonuclease A. *Chemical Reviews* **98,** 1045–1066 (1998).

76. Ivarsson, Y. Plasticity of PDZ domains in ligand recognition and signaling. *FEBS Letters* **586,** 2638–2647 (2012).

77. Van Ham, M. & Hendriks, W. PDZ domains-glue and guide. *Mol. Biol. Rep.* **30,** 69–82 (2003).

78. Spaller, M. R. Act Globally, Think Locally: Systems Biology Addresses the PDZ Domain. *ACS Chemical Biology* **1,** 207–210 (2006).

79. Garner, C. C., Nash, J. & Huganir, R. L. PDZ domains in synapse assembly and signalling. *Trends in Cell Biology* **10,** 274–280 (2000).

80. Sheng, M. & Sala, C. PDZ Domains and the Organization of Supramolecular Complexes. *Annual Review of Neuroscience* **24,** 1–29 (2001).

81. Zhang, M. & Wang, W. Organization of Signaling Complexes by PDZ-Domain Scaffold Proteins. *Accounts of Chemical Research* **36,** 530–538 (2003).

82. Roh, M. H. & Margolis, B. Composition and function of PDZ protein complexes during cell polarization. *American Journal of Physiology-Renal Physiology* **285,** F377–F387 (2003).

83. Fan, J.-S. & Zhang, M. Signaling Complex Organization by PDZ Domain Proteins. *Neurosignals* **11,** 315–321 (2002).

84. Craven, S. E. & Bredt, D. S. PDZ Proteins Organize Synaptic Signaling Pathways. *Cell* **93,** 495–498 (1998).

85. Brenman, J. E. *et al.* Localization of Postsynaptic Density-93 to Dendritic Microtubules and Interaction with Microtubule-Associated Protein 1A. *Journal of Neuroscience* **18,** 8805–8813 (1998).

86. Gomperts, S. N. Clustering Membrane Proteins: It's All Coming Together with the PSD-95/SAP90 Protein Family. *Cell* **84,** 659–662 (1996).

87. Schultz, J., Milpetz, F., Bork, P. & Ponting, C. P. SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences* **95,** 5857–5864 (1998).

88. Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P. & Bork, P. SMART: a web-based tool for the study of genetically mobile domains. *Nucleic acids research* **28,** 231–234 (2000).

89. Ponting, C. P. Evidence for PDZ domains in bacteria, yeast, and plants. *Protein Science* **6,** 464–468 (1997).

90. Pallen, M. J. & Ponting, C. P. PDZ domains in bacterial proteins. *Molecular Microbiology* **26,** 411–413 (1997).

91. Doyle, D. A. *et al.* Crystal Structures of a Complexed and Peptide-Free Membrane Protein-Binding Domain: Molecular Basis of Peptide Recognition by PDZ. *Cell* **85,** 1067–1076 (1996).

92. Fanning, A. S. & Anderson, J. M. Protein-protein interactions: PDZ domain networks. *Current Biology* **6,** 1385–1388 (1996).

93. Harrison, S. C. Peptide-Surface Association: The Case of PDZ and PTB Domains. *Cell* **86,** 341–343 (1996).

94. Remaut, H. & Waksman, G. Protein-protein interaction through β-strand addition. *Trends in Biochemical Sciences* **31,** 436–444 (2006).

95. Songyang, Z. *et al.* Recognition of Unique Carboxyl-Terminal Motifs by Distinct PDZ Domains. *Science* **275,** 73–77 (1997).

96. Schultz, J. *et al.* Specific interactions between the syntrophin PDZ domain and voltage-gated sodium channels. *Nature Structural Biology* **5,** 19–24 (1998).

97. Daniels, D. L., Cohen, A. R., Anderson, J. M. & Brünger, A. T. Crystal structure of the hCASK PDZ domain reveals the structural basis of class II PDZ domain target recognition. *Nature Structural Biology* **5,** 317–325 (1998).

98. Hillier, B. J., Christopherson, K. S., Prehoda, K. E., Bredt, D. S. & Lim, W. A. Unexpected Modes of PDZ Domain Scaffolding Revealed by Structure of nNOS-Syntrophin Complex. *Science* **284,** 812–815 (1999).

99. Kiick, K. L., Saxon, E., Tirrell, D. A. & Bertozzi, C. R. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 19–24 (2002).

100. Simon, M., Zangemeister-Wittke, U. & Plückthun, A. Facile double functionalization of designed ankyrin repeat proteins using click and thiol chemistries. *Bioconjugate Chem.* **23,** 279–286 (2012).

101. Stucki-Buchli, B. *et al.* 2D-IR Spectroscopy of an AHA Labeled Photoswitchable PDZ2 Domain. *The Journal of Physical Chemistry A* **121,** 9435–9445 (2017).

102. Bloem, R. *et al.* Ligand Binding Studied by 2D IR Spectroscopy Using the Azidohomoalanine Label. *The Journal of Physical Chemistry B* **116,** 13705–13712 (2012).

103. Johnson, P. J. M., Koziol, K. L. & Hamm, P. Quantifying Biomolecular Recognition with Site-Specific 2D Infrared Probes. *The Journal of Physical Chemistry Letters* **8,** 2280–2284 (2017).

104. Chan, W. C. & White, P. D. in *Fmoc solid phase peptide synthesis: a practical approach* (New York: Oxford University Press, 2000).

105. Wachtveitl, J. & Zumbusch, A. Azobenzene: An Optical Switch for in vivo Experiments. *ChemBioChem* **12,** 1169–1170 (2011).

106. Garcia-Amorós, J., Díaz-Lobo, M., Nonell, S. & Velasco, D. Fastest Thermal Isomerization of an Azobenzene for Nanosecond Photoswitching Applications under Physiological Conditions. *Angewandte Chemie International Edition* **51,** 12820–12823 (2012).

107. Goeldner, M. & Givens, R. Dynamic Studies in Biology: Phototriggers, Photoswitches and Caged Biomolecules (2005).

108. Ellis-Davies & Graham, C. R. Caged compounds: photorelease technology for control of cellular chemistry and physiology. *Nature Methods* **4,** 619–628 (Aug. 2007).

109. Lee, H.-M., Larson, D. R. & Lawrence, D. S. Illuminating the Chemistry of Life: Design, Synthesis, and Applications of “Caged” and Related Photoresponsive Compounds. *ACS Chemical Biology* **4,** 409–427 (2009).

110. Noble, A. III. Zur Geschichte des Azobenzols und des Benzidins. *Justus Liebigs Annalen der Chemie* **98,** 253–256 (1856).

111. Griffiths, J. II. Photochemistry of azobenzene and its derivatives. *Chem. Soc. Rev.* **1,** 481–493 (4 1972).

112. Dias, A. *et al.* Enthalpies of formation of cis-azobenzene and trans-azobenzene. *The Journal of Chemical Thermodynamics* **24,** 439–447 (1992).

113. Bandara, H. M. D. & Burdette, S. C. Photoisomerization in different classes of azobenzene. *Chem. Soc. Rev.* **41,** 1809–1825 (5 2012).

114. Fischer, E., Frankel, M. & Wolovsky, R. Wavelength Dependence of Photoisomerization Equilibria in Azocompounds. *The Journal of Chemical Physics* **23,** 1367–1367 (1955).

115. Fischer, E. Temperature Dependence of Photoisomerization Equilibria. Part I. Azobenzene and the Azonaphthalenes. *Journal of the American Chemical Society* **82,** 3249–3252 (1960).

116. Bortolus, P. & Monti, S. Cis-trans photoisomerization of azobenzene. Solvent and triplet donors effects. *The Journal of Physical Chemistry* **83,** 648–652 (1979).

117. Wang, L., Xu, W., Yi, C. & Wang, X. Isomerization and electronic relaxation of azobenzene after being excited to higher electronic states. *Journal of Molecular Graphics and Modelling* **27,** 792–796 (2009).

118. Cattaneo, P. & Persico, M. An abinitio study of the photochemistry of azobenzene. *Phys. Chem. Chem. Phys.* **1,** 4739–4743 (1999).

119. Granucci, G. & Persico, M. Excited state dynamics with the direct trajectory surface hopping method: azobenzene and its derivatives as a case study. *Theoretical Chemistry Accounts* **117,** 1131–1143 (May 2007).

120. Rau, H. in *Photoreactive Organic Thin Films* (eds Sekkat, Z. & Knoll, W.) 3–47 (Academic Press, San Diego, 2002).

121. Siampiringue, N., Guyot, G., Monti, S. & Bortolus, P. The cis →trans photoisomerization of azobenzene: an experimental re-examination. *Journal of Photochemistry* **37,** 185–188 (1987).

122. Rau, H. & Lueddecke, E. On the rotation-inversion controversy on photoisomerization of azobenzenes. Experimental proof of inversion. *Journal of the American Chemical Society* **104,** 1616–1620 (1982).

123. Monti, S., Orlandi, G. & Palmieri, P. Features of the photochemically active state surfaces of azobenzene. *Chemical Physics* **71,** 87–99 (1982).

124. Crecca, C. R. & Roitberg, A. E. Theoretical Study of the Isomerization Mechanism of Azobenzene and Disubstituted Azobenzene Derivatives. *The Journal of Physical Chemistry A* **110,** 8188–8203 (2006).

125. Chang, C.-W., Lu, Y.-C., Wang, T.-T. & Diau, E. W.-G. Photoisomerization Dynamics of Azobenzene in Solution with S1 Excitation: A Femtosecond Fluorescence Anisotropy Study. *Journal of the American Chemical Society* **126,** 10109–10118 (2004).

126. Wang, L., Xu, W., Yi, C. & Wang, X. Isomerization and electronic relaxation of azobenzene after being excited to higher electronic states. *Journal of Molecular Graphics and Modelling* **27,** 792–796 (2009).

127. Schultz, T. *et al.* Mechanism and Dynamics of Azobenzene Photoisomerization. *Journal of the American Chemical Society* **125,** 8098–8099 (2003).

128. Samanta, S. *et al.* Photoswitching Azo Compounds in Vivo with Red Light. *Journal of the American Chemical Society* **135,** 9777–9784 (2013).

129. Dong, M., Babalhavaeji, A., Samanta, S., Beharry, A. A. & Woolley, G. A. Red-Shifting Azobenzene Photoswitches for in Vivo Use. *Accounts of Chemical Research* **48,** 2662–2670 (2015).

130. Sadovski, O., Beharry, A. A., Zhang, F. & Woolley, G. A. Spectral Tuning of Azobenzene Photoswitches for Biological Applications. *Angewandte Chemie* **121,** 1512–1514 (2009).

131. Zhang, Z., Burns, D. C., Kumita, J. R., Smart, O. S. & Woolley, G. A. A Water-Soluble Azobenzene Cross-Linker for Photocontrol of Peptide Conformation. *Bioconjugate Chemistry* **14,** 824–829 (2003).

132. Chi, L., Sadovski, O. & Woolley, G. A. A Blue-Green Absorbing Cross-Linker for Rapid Photoswitching of Peptide Helix Content. *Bioconjugate Chemistry* **17,** 670–676 (2006).

133. Beharry, A. A., Sadovski, O. & Woolley, G. A. Azobenzene Photoswitching without Ultraviolet Light. *Journal of the American Chemical Society* **133,** 19684–19687 (2011).

134. Burns, D. C., Zhang, F. & Woolley, G. A. Synthesis of 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido)azobenzene and cysteine cross-linking for photo-control of protein conformation and activity. *Nat. Protoc.* **2,** 251–258 (2007).

135. Bredenbeck, J., Helbing, J., Kumita, J. R., Woolley, G. A. & Hamm, P. α-Helix formation in a photoswitchable peptide tracked from picoseconds to microseconds by time-resolved IR spectroscopy. *Proceedings of the National Academy of Sciences* **102,** 2379–2384 (2005).

136. Bredenbeck, J. *et al.* Picosecond conformational transition and equilibration of a cyclic peptide. *Proceedings of the National Academy of Sciences* **100,** 6452–6457 (2003).

137. Renner, C., Kusebauch, U., Löweneck, M., Milbradt, A. & Moroder, L. Azobenzene as photoresponsive conformational switch in cyclic peptides*. *The Journal of Peptide Research* **65,** 4–14 (2005).

138. Renner, C., Behrendt, R., Spörlein, S., Wachtveitl, J. & Moroder, L. Photomodulation of conformational states. I. Mono- and bicyclic peptides with (4-amino)phenylazobenzoic acid as backbone constituent. *Biopolymers* **54,** 489–500 (2000).

139. Wachtveitl, J. *et al.* Ultrafast Conformational Dynamics in Cyclic Azobenzene Peptides of Increased Flexibility. *Biophysical Journal* **86,** 2350–2362 (2004).

140. Satzger, H. *et al.* Picosecond dynamics in water-soluble azobenzene-peptides. *Chemical Physics Letters* **396,** 191–197 (2004).

141. Borisenko, V. & Woolley, G. A. Reversibility of conformational switching in light-sensitive peptides. *Journal of Photochemistry and Photobiology A: Chemistry* **173,** 21–28 (2005).

142. Woolley, G. A. Photocontrolling Peptide α Helices. *Accounts of Chemical Research* **38,** 486–493 (2005).

143. Flint, D. G., Kumita, J. R., Smart, O. S. & Woolley, G. Using an Azobenzene Cross-Linker to Either Increase or Decrease Peptide Helix Content upon Trans-to-Cis Photoisomerization. *Chemistry Biology* **9,** 391–397 (2002).

144. Kumita, J. R., Smart, O. S. & Woolley, G. A. Photo-control of helix content in a short peptide. *Proceedings of the National Academy of Sciences* **97,** 3803–3808 (2000).

145. Buchli, B. *et al.* Kinetic response of a photoperturbed allosteric protein. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 11725–11730 (2013).

146. Waldauer, S. A., Stucki-Buchli, B., Frey, L. & Hamm, P. Effect of viscogens on the kinetic response of a photoperturbed allosteric protein. *J. Chem. Phys.* **141,** 22D514 (2014).

147. Poole, L. B. The basics of thiols and cysteines in redox biology and chemistry. *Free Radical Biology and Medicine* **80,** 148–157 (2015).

148. Jankovic, B. *et al.* Photocontrolling Protein-Peptide Interactions: From Minimal Perturbation to Complete Unbinding. *Journal of the American Chemical Society* **141,** 10702–10710 (2019).

149. Waldauer, S. A. *et al.* Photocontrol of Reversible Amyloid Formation with a Minimal-Design Peptide. *The Journal of Physical Chemistry B* **116,** 8961–8973 (2012).

150. Marshall, G. R., Feng, J. A. & Kuster, D. J. Back to the future: Ribonuclease A. *Peptide Science* **90,** 259–277 (2008).

151. Cuchillo, C. M., Nogués, M. V. & Raines, R. T. Bovine Pancreatic Ribonuclease: Fifty Years of the First Enzymatic Reaction Mechanism. *Biochemistry* **50,** 7835–7841 (2011).

152. Anfinsen, C. B. Principles that Govern the Folding of Protein Chains. *Science* **181,** 223–230 (1973).

153. Moore, S. & Stein, W. H. Chemical Structures of Pancreatic Ribonuclease and Deoxyribonuclease. *Science* **180,** 458–464 (1973).

154. Richards, F. M. The 1972 Nobel Prize for Chemistry. *Science* **178,** 492–493 (1972).

155. Merrifield, B. Solid phase synthesis. *Science* **232,** 341–347 (1986).

156. Kaiser, E. T. The 1984 Nobel Prize in Chemistry. *Science* **226,** 1151–1153 (1984).

157. Barnard, E. A. Ribonucleases. *Annual Review of Biochemistry* **38,** 677–732 (1969).

158. Cech, T. R. The RNA worlds in context. *Cold Spring Harbor perspectives in biology* **4(7)** (2012).

159. Barnard, E. A. Biological Function of Pancreatic Ribonuclease. *Nature* **221,** 340–344 (1969).

160. Kunitz, M. CRYSTALLINE RIBONUCLEASE. *The Journal of General Physiology* **24,** 15–32 (1940).

161. McDonald, M. R. A method for the preparation of protease-free crystalline ribonuclease. *The Journal of general physiology* **32,** 39–42 (1948).

162. Crestfield, A. M., Stein, W. H. & Moore, S. F. On the preparation of bovine pancreatic ribonuclease A. *The Journal of biological chemistry* **238,** 618–621 (1963).

163. Klink, T. A., Woycechowsky, K. J., Taylor, K. M. & Raines, R. T. Contribution of disulfide bonds to the conformational stability and catalytic activity of ribonuclease A. *European Journal of Biochemistry* **267,** 566–572 (2000).

164. Neira, J. L. & Rico, M. Folding studies on ribonuclease A, a model protein. *Folding and Design* **2,** R1–R11 (1997).

165. Schultz, D. A., Baldwin, R. L. & Schmid, F. X. Cis proline mutants of ribonuclease A. II. Elimination of the slow-folding forms by mutation. *Protein Science* **1,** 917–924 (1992).

166. Thompson, J. E. & Raines, R. T. Value of General Acid-Base Catalysis to Ribonuclease A. *Journal of the American Chemical Society* **116,** 5467–5468 (1994).

167. Deakyne, C. A. & Allen, L. C. Role of active-site residues in the catalytic mechanism of ribonuclease A. *Journal of the American Chemical Society* **101,** 3951–3959 (1979).

168. Findlay, D., Herries, D. G., Mathias, A. P., Rabin, B. R. & Ross, C. A. The Active Site and Mechanism of Action of Bovine Pancreatic Ribonuclease. *Nature* **190,** 781–784 (1961).

169. Wlodawer, A., Svensson, L. A., Sjoelin, L. & Gilliland, G. L. Structure of phosphate-free ribonuclease A refined at 1.26 .ANG. *Biochemistry* **27,** 2705–2717 (1988).

170. Richards, F. M. & Vithayathil, P. J. The Preparation of Subtilisin-modified Ribonuclease and the Separation of the Peptide and Protein Components. *Journal of Biological Chemistry* **234,** 1459–1465 (1959).

171. Kim, E. E., Varadarajan, R., Wyckoff, H. W. & Richards, F. M. Refinement of the crystal structure of Ribonuclease S. Comparison with and between the various Ribonuclease A structures. *Biochemistry* **31,** 12304–12314 (1992).

172. Wlodawer, A. & Sjölin, L. Hydrogen exchange in RNase A: neutron diffraction study. *Proceedings of the National Academy of Sciences* **79,** 1418–1422 (1982).

173. Luitz, M. P., Bomblies, R. & Zacharias, M. Comparative Molecular Dynamics Analysis of RNase-S Complex Formation. *Biophysical journal* **113,** 1466–1474 (2017).

174. Cotesta, S., Tavernelli, I. & Iorio, E. E. D. Dynamics of RNase-A and S-Protein: A Molecular Dynamics Simulation of the Transition Toward a Folding Intermediate. *Biophysical Journal* **85,** 2633–2640 (2003).

175. Bachmann, A., Wildemann, D., Praetorius, F., Fischer, G. & Kiefhaber, T. Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction. *Proceedings of the National Academy of Sciences* **108,** 3952–3957 (2011).

176. Goldberg, J. M. & Baldwin, R. L. A specific transition state for S-peptide combining with folded S-protein and then refolding. *Proceedings of the National Academy of Sciences* **96,** 2019–2024 (1999).

177. Schreier, A. A. & Baldwin, R. L. Mechanism of dissociation of S-peptide from Ribonuclease S. *Biochemistry* **16,** 4203–4209 (1977).

178. Changeux, J.-P. & Edelstein, S. Conformational selection or induced fit? 50 years of debate resolved. *F1000 biology reports* **3** (2011).

179. Dogan, J., Gianni, S. & Jemth, P. The binding mechanisms of intrinsically disordered proteins. *Phys. Chem. Chem. Phys.* **16,** 6323–6331 (14 2014).

180. Wlodawer, A. Comparison of Two Independently Refined Models of Ribonuclease-A. *Acta Crystallographica* **B42,** 379–387 (1986).

181. Kozlov, G., Banville, D., Gehring, K. & Ekiel, I. Solution Structure of the PDZ2 Domain from Cytosolic Human Phosphatase hPTP1E Complexed with a Peptide Reveals Contribution of the β2-β3 Loop to PDZ Domain-Ligand Interactions. *Journal of Molecular Biology* **320,** 813–820 (2002).

182. Blöchliger, N., Xu, M. & Caflisch, A. Peptide Binding to a PDZ Domain by Electrostatic Steering via Nonnative Salt Bridges. *Biophysical Journal* **108,** 2362–2370 (2015).

183. Everything you wanted to know about Markov State Models but were afraid to ask. *Methods* **52,** 99–105 (2010).

184. Prinz, J.-H. *et al.* Markov models of molecular kinetics: Generation and validation. *The Journal of Chemical Physics* **134,** 174105 (2011).

185. Chodera, J. D. & Noé, F. Markov state models of biomolecular conformational dynamics. *Current Opinion in Structural Biology* **25.** Theory and simulation / Macromolecular machines, 135–144 (2014).

186. Sengupta, U. & Strodel, B. Markov models for the elucidation of allosteric regulation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **373,** 20170178 (2018).

187. Bowman, G. R., Beauchamp, K. A., Boxer, G. & Pande, V. S. Progress and challenges in the automated construction of Markov state models for full protein systems. *The Journal of Chemical Physics* **131,** 124101 (2009).

188. Schwantes, C. R. & Pande, V. S. Improvements in Markov State Model Construction Reveal Many Non-Native Interactions in the Folding of NTL9. *Journal of Chemical Theory and Computation* **9,** 2000–2009 (2013).

189. Cooper, A. & Dryden, D. T. F. Allostery without conformational change. A plausible model. *Eur. Biophys. J.* **11,** 103–109 (1984).

190. Cavalli, A., Salvatella, X., Dobson, C. M. & Vendruscolo, M. Protein structure determination from NMR chemical shifts. *Proceedings of the National Academy of Sciences* **104,** 9615–9620 (2007).

191. Stock, G. & Hamm, P. A non-equilibrium approach to allosteric communication. *Philosophical Transactions of the Royal Society B: Biological Sciences* **373,** 20170187 (2018).

192. Laskowski, R. A., Gerick, F. & Thornton, J. M. The structural basis of allosteric regulation in proteins. *FEBS Letters* **583,** 1692–1698 (2009).

193. Toto, A., Mattei, A., Jemth, P. & Gianni, S. Understanding the role of phosphorylation in the binding mechanism of a PDZ domain. *Protein Engineering, Design and Selection* **30,** 1–5 (Dec. 2016).

194. Cao, T. T., Deacon, H. W., Reczek, D., Bretscher, A. & von Zastrow, M. A kinase-regulated PDZ-domain interaction controls endocytic sorting of the β2-adrenergic receptor. *Nature* **401,** 286–290 (1999).

195. Gianni, S. *et al.* The Kinetics of PDZ Domain-Ligand Interactions and Implications for the Binding Mechanism. *Journal of Biological Chemistry* **280,** 34805–34812 (2005).

196. Krupka, R. M., Kaplan, H. & Laidler, K. J. Kinetic consequences of the principle of microscopic reversibility. *Trans. Faraday Soc.* **62,** 2754–2759 (1966).

197. Bhatt, D. & Zuckerman, D. M. Beyond Microscopic Reversibility: Are Observable Nonequilibrium Processes Precisely Reversible? *Journal of Chemical Theory and Computation* **7,** 2520–2527 (2011).

198. Poirier, M., Eroglu, S., Chatenay, D. & Marko, J. F. Reversible and Irreversible Unfolding of Mitotic Newt Chromosomes by Applied Force. *Molecular Biology of the Cell* **11,** 269–276 (2000).

199. Andreev, O. A. *et al.* pH (low) insertion peptide (pHLIP) inserts across a lipid bilayer as a helix and exits by a different path. *Proceedings of the National Academy of Sciences* **107,** 4081–4086 (2010).

200. Lórenz-Fonfría, V. A. & Kandori, H. Transformation of Time-Resolved Spectra to Lifetime-Resolved Spectra by Maximum Entropy Inversion of the Laplace Transform. *Appl. Spectrosc.* **60,** 407–417 (Apr. 2006).

201. Kumar, A. T. N., Zhu, L., Christian, J. F., Demidov, A. A. & Champion, P. M. On the Rate Distribution Analysis of Kinetic Data Using the Maximum Entropy Method: Applications to Myoglobin Relaxation on the Nanosecond and Femtosecond Timescales. *The Journal of Physical Chemistry B* **105,** 7847–7856 (2001).

202. Kim, J.-S. & Raines, R. T. Ribonuclease S-peptide as a carrier in fusion proteins. *Protein Science* **2,** 348–356 (1993).

203. Schreiber, G. & Keating, A. E. Protein binding specificity versus promiscuity. *Current Opinion in Structural Biology* **21,** 50–61 (2011).

204. Mendel, C. M. & Mendel, D. B. 'Non-specific' binding. The problem, and a solution. *Biochemical Journal* **228,** 269–272 (May 1985).

205. Law, A. B., Fuentes, E. J. & Lee, A. L. Conservation of side-chain dynamics within a protein family. *Journal of the American Chemical Society* **131,** 6322–6323 (May 2009).

206. Cowgill, R. W. Fluorescence and the structure of proteins: III. Effects of denaturation on fluorescence of insulin and ribonuclease. *Archives of Biochemistry and Biophysics* **104,** 84–92 (1964).

207. Noronha, M., Lima, J. C., Paci, E., Santos, H. & Maçanita, A. L. Tracking Local Conformational Changes of Ribonuclease A Using Picosecond Time-Resolved Fluorescence of the Six Tyrosine Residues. *Biophysical Journal* **92,** 4401–4414 (2007).

208. Petit, C. M., Zhang, J., Sapienza, P. J., Fuentes, E. J. & Lee, A. L. Hidden dynamic allostery in a PDZ domain. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 18249–18254 (2009).

209. Saro, D. *et al.* A Thermodynamic Ligand Binding Study of the Third PDZ Domain (PDZ3) from the Mammalian Neuronal Protein PSD-95. *Biochemistry* **46,** 6340–6352 (2007).

210. Steiner, S. & Caflisch, A. Peptide binding to the PDZ3 domain by conformational selection. *Proteins: Structure, Function, and Bioinformatics* **80,** 2562–2572 (2012).

211. Zhang, J., Petit, C. M., King, D. S. & Lee, A. L. Phosphorylation of a PDZ Domain Extension Modulates Binding Affinity and Interdomain Interactions in Postsynaptic Density-95 (PSD-95) Protein, a Membrane-associated Guanylate Kinase (MAGUK). *Journal of Biological Chemistry* **286,** 41776–41785 (2011).

212. Bredenbeck, J., Helbing, J. & P.Hamm. Continuous scanning from picoseconds to microseconds in time resolved linear and nonlinear spectroscopy. *Review of Scientific Instruments* **75,** 4462–4466 (2004).

213. Hamm, P., Kaindl, R. A. & Stenger, J. Noise suppression in femtosecond mid-infrared light sources. *Opt. Lett.* **25,** 1798–1800 (Dec. 2000).

214. Feng, Y., Vinogradov, I. & Ge, N.-H. General noise suppression scheme with reference detection in heterodyne nonlinear spectroscopy. *Opt. Express* **25,** 26262–26279 (Oct. 2017).

215. Bertero, M. & Boccacci, P. *Introduction to inverse problems in imaging* (CRC press, 1998).

216. Lavrentiev, M., Romanov, V. & Shishatskii, S. *Ill-posed Problems of Mathematical Physics and Analysis* (American Mathematical Society, 1986).

217. Engl, H. W. & Groetsch, C. *Inverse and Ill-Posed Problems* (Academic Press, 1987).

218. Hansen, P. C. & O'Leary, D. P. The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems. *SIAM Journal on Scientific Computing* **14,** 1487–1503 (1993).

219. Tikhonov, A. & Arsenin, V. *Solutions of ill-posed problems* (Winston, 1977).

220. Skilling, J. & Gull, S. F. in *Maximum-Entropy and Bayesian Methods in Inverse Problems* (eds Smith, C. R. & Grandy, W. T.) 83–132 (Springer Netherlands, Dordrecht, 1985).

221. Skilling, J. & Bryan, R. K. Maximum entropy image reconstruction: general algorithm. *Monthly Notices of the Royal Astronomical Society* **211,** 111–124 (Nov. 1984).

222. Jaynes, E. T. Information Theory and Statistical Mechanics. *Phys. Rev.* **106,** 620–630 (4 May 1957).

223. Chiang, Y.-W., Borbat, P. P. & Freed, J. H. Maximum entropy: A complement to Tikhonov regularization for determination of pair distance distributions by pulsed ESR. *Journal of Magnetic Resonance* **177,** 184–196 (2005).

224. Lesne, A. Shannon entropy: a rigorous notion at the crossroads between probability, information theory, dynamical systems and statistical physics. *Mathematical Structures in Computer Science* **24,** e240311 (2014).

225. Shannon, C. E. A Mathematical Theory of Communication. *Bell System Technical Journal* **27,** 379–423 (1948).

226. Gull, S. & Daniell, G. Image reconstruction from incomplete and noisy data. *Nature* **272,** 686–690 (Apr. 1978).

227. Sibisi, S., Skilling, J., Brereton, R. G., Laue, E. D. & Staunton, J. Maximum entropy signal processing in practical NMR spectroscopy. *Nature* **311,** 446–447 (Oct. 1984).

228. Brochon, J.-C. in *Part B: Numerical Computer Methods* 262–311 (Academic Press, 1994).

229. Hobson, M. P. & Lasenby, A. N. The entropic prior for distributions with positive and negative values. *Monthly Notices of the Royal Astronomical Society* **298,** 905–908 (Aug. 1998).

230. P. J. M. Johnson, K. L. Koziol, and P. Hamm. Intrinsic phasing of heterodyne-detected multidimensional infrared spectra. *Opt. Express* **25,** 2928–2938 (Feb. 2017).

231. Bloem, R., Garrett-Roe, S., Strzalka, H., Hamm, P. & Donaldson, P. Enhancing signal detection and completely eliminating scattering using quasi-phase-cycling in 2D IR experiments. *Opt. Express* **18,** 27067–27078 (Dec. 2010).

232. Hamm, P. & Zanni, M. in *Concepts and Methods of 2D Infrared Spectroscopy* 88–108 (Cambridge University Press, 2011).

233. Ouseph, P. J., Driver, K. & Conklin, J. Polarization of light by reflection and the Brewster angle. *American Journal of Physics* **69,** 1166–1168 (2001).

234. Hamm, P. & Zanni, M. in *Concepts and Methods of 2D Infrared Spectroscopy* 202 (Cambridge University Press, 2011).

235. Hamm, P. & Zanni, M. in *Concepts and Methods of 2D Infrared Spectroscopy* 61–87 (Cambridge University Press, 2011).

236. Volkov, V., Schanz, R. & Hamm, P. Active phase stabilization in Fourier-transform two-dimensional infrared spectroscopy. *Opt. Lett.* **30,** 2010–2012 (Aug. 2005).

237. Hamm, P. & Zanni, M. in *Concepts and Methods of 2D Infrared Spectroscopy* 176–216 (Cambridge University Press, 2011).

238. Backus, E. H. G., Garrett-Roe, S. & Hamm, P. Phasing problem of heterodyne-detected two-dimensional infrared spectroscopy. *Opt. Lett.* **33,** 2665–2667 (Nov. 2008).

239. Venyaminov, S. Y. & Prendergast, F. G. Water (H_2O and D_2O) Molar Absorptivity in the 100-4000 cm^{-1} Range and Quantitative Infrared Spectroscopy of Aqueous Solutions. *Analytical Biochemistry* **248,** 234–245 (1997).

240. Piatkowski, L., Eisenthal, K. B. & Bakker, H. J. Ultrafast intermolecular energy transfer in heavy water. *Phys. Chem. Chem. Phys.* **11,** 9033–9038 (2009).

Dr. Claudio Zanobini

Eibenstrasse 18/89, 8045 Zurich
| +41 76 379 5980 **|** claudio.zanobini@chem.uzh.ch

Profile

Date of birth	13th February 1987
Languages	Italian *(native proficiency)*, English *(full professional proficiency),* German *(intermediate),* Spanish *(advanced),* Brazilian Portuguese *(basic/intermediate)*
Software	LabVIEW, Python (with certificate), MS Office, Origin, LaTeX

Experience

2/2015 – 6/2020 **PhD in Chemistry/Molecular Spectroscopy** *(University of Zurich, CH)*

- Ultrafast (femtosecond) laser spectroscopy
- Optics, Photonics
- Two-Dimensional Infrared, Transient Infrared Pump-Probe Spectroscopies
- Hands-on laboratory testing, optical alignment and measurements
- Protein Biophysics/Proteomics/Biochemistry
- Bioanalytical Chemistry/protein separation techniques, chromatography and electrophoresis
- Experience in scientific programming using Python, and LabVIEW

2/2015 – 6/2017 **Teaching Assistant** *(University of Zurich, CH)*

- Person in charge of the assistance of 20-25 students

Education

09/2011– 12/2014 **MSc in Chemistry, grade: 110/110** *(University of Florence, IT)*

- Molecular Spectroscopy
- Physical Chemistry

09/2006– 4/2011 **BSc in Chemistry, grade: 110/110** *(University of Florence, IT)*

- Chemistry, Physics, Maths

09/2001– 7/2006 **Higher Education Diploma, grade: 100/100** *(Giorgio Vasari Inst., IT)*

- Technician Qualified in the Management of the Administering of Foods and Beverages

List of Publications

[1] Brankica Jankovic, Adnan Gulzar, Claudio Zanobini, Olga Bozovic, Steffen Wolf, Gerhard Stock, Peter Hamm, “Photocontrolling protein-peptide interactions: From minimal perturbation to complete unbinding”, Journal of the American Chemical Society, vol. 141, no. 27, pp. 10 702–10 710, 2019.

[2] Claudio Zanobini, Olga Bozovic, Brankica Jankovic, Klemens L. Koziol, Philip J. M. Johnson, Peter Hamm, Adnan Gulzar, Steffen Wolf, Gerhard Stock, “Azidohomoalanine: A Minimally Invasive, Versatile, and Sensitive Infrared Label in Proteins To Study Ligand Binding”, The Journal of Physical Chemistry B, vol. 122, no. 44, pp. 10 118–10 125, Oct. 2018.

[3] Brigitte Stucki-Buchli, Philip J. M. Johnson, Olga Bozovic, Claudio Zanobini, Klemens L. Koziol, Peter Hamm, Adnan Gulzar, Steffen Wolf, Sebastian Buchenberg, Gerhard Stock, “2D-IR Spectroscopy of an AHA Labeled Photoswitchable PDZ2 Domain”, The Journal of Physical Chemistry A, vol. 121, no. 49, pp. 9435–9445, 2017.


CMSZH
CHEMICAL MOLECULAR SCIENCES ZURICH
GRADUATE SCHOOL